4-2024

My GitHub Sponsors profile is live!": Investigating the impact of Twitter/X mentions on GitHub Sponsors

Youmei FAN

Tao XIAO

Hideaki HATA

Christoph TREUDE
Singapore Management University, ctreude@smu.edu.sg

Kenichi MATSUMOTO

Follow this and additional works at: https://ink.library.smu.edu.sg/sis_research

Part of the Social Media Commons, and the Software Engineering Commons

"My GitHub Sponsors profile is live!" Investigating the Impact of Twitter/X Mentions on GitHub Sponsors

Youmei Fan
Nara Institute of Science and
Technology, Japan
fan.youmei.fs2@is.naist.jp

Tao Xiao
Nara Institute of Science and
Technology, Japan
tao.xiao.ts2@is.naist.jp

Hideaki Hata
Shinshu University
Japan
hata@shinshu-u.ac.jp

Christoph Treude
University of Melbourne
Australia
christoph.treude@unimelb.edu.au

Kenichi Matsumoto
Nara Institute of Science and
Technology, Japan
matumoto@is.naist.jp

ABSTRACT

GitHub Sponsors was launched in 2019, enabling donations to open-source software developers to provide financial support, as per GitHub's slogan: "Invest in the projects you depend on". However, a 2022 study on GitHub Sponsors found that only two-fifths of developers who were seeking sponsorship received a donation. The study found that, other than internal actions (such as offering perks to sponsors), developers had advertised their GitHub Sponsors profiles on social media, such as Twitter (also known as X). Therefore, in this work, we investigate the impact of tweets that contain links to GitHub Sponsors profiles on sponsorship, as well as their reception on Twitter/X. We further characterize these tweets to understand their context and find that (1) such tweets have the impact of increasing the number of sponsors acquired, (2) compared to other donation platforms such as Open Collective and Patreon, GitHub Sponsors has significantly fewer interactions but is more visible on Twitter/X, and (3) developers tend to contribute more to open-source software during the week of posting such tweets. Our findings are the first step toward investigating the impact of social media on obtaining funding to sustain open-source software.

CCS CONCEPTS

• **Social and professional topics** → **Sustainability**; • **Software and its engineering** → **Open source model**.

KEYWORDS

Open-source Software, Sponsorship, Social Media

1 INTRODUCTION

Open-source software (OSS) is ubiquitous, but sustaining it is a challenge [28]. Maintaining an OSS project requires not only intrinsic motivation (e.g., joy of participation) but also extrinsic motivation (e.g., financial incentives) [49]. The last few years have seen the emergence of many platforms that allow open-source developers to receive donations for their work, such as PayPal [30], Open Collective [11], Patreon [29], and GitHub Sponsors [37]. Several platforms support sponsoring OSS projects in cryptocurrencies, with the rise in popularity of cryptocurrencies today, e.g., Gitcoin Grants [16] and Giveth [20]. However, as Overney et al.'s paper title "How to not get rich: an empirical study of donations in open source" [28] suggests, simply having a platform for donations is not enough. In a 2022 study on GitHub Sponsors, Shimada et al. [35] found that out of approximately 9,000 developers who had activated their GitHub Sponsors profile, less than 40% had received a donation.

If simply creating a sponsorship profile is not enough, what else can open-source software developers do to attract donations? Following the long line of work on studying the intersection between social media and software development [15, 38, 39], in this paper, we investigate the impact of tweeting about a GitHub Sponsors profile on sponsorship. To make it easy for its users to reach a large audience, GitHub provides tweet templates that users can use to advertise a new GitHub Sponsors profile ("My GitHub Sponsors profile is live! You can sponsor me to support my open source work 💕") or to broadcast that they made a donation ("💕 I'm sponsoring [username] because..."). The impact of tweets on open-source software development has been investigated before. Fang et al. [15] found that tweets have a significant effect on obtaining new stars and new contributors for an open-source project and that the formation of an active Twitter/X community plays an important role in attracting new contributors. The role of tweets has also been studied in the context of bug fixing [24] and trend awareness [36]. To the best of our knowledge, the role of Twitter/X in obtaining funding for open-source development has not yet been studied.

We first characterize the state of the practice by quantitatively and qualitatively analyzing more than 10,000 tweets linking to GitHub Sponsors profiles to understand the context of such tweets.

We then measure the impact of the tweets in terms of their reception on Twitter/X and their effect on sponsorship. We found links to GitHub Sponsors profiles on Twitter/X are common and the majority of such tweets are written by users other than the profile owner, such as a sponsor. We identified a significant positive effect of GitHub Sponsor profile mentions in tweets on the number of sponsors acquired. Tweet mentions have the impact of increasing the number of sponsors by 1.22. Although GitHub Sponsors has surpassed other platforms such as Open Collective and Patreon in terms of visibility on Twitter/X, tweets about GitHub Sponsors received significantly fewer likes, retweets, and replies compared to other platforms. Developers tended to be more active during the week of a tweet, in particular, in terms of the number of commits.

Significance of research contribution. The findings of our study have significant implications, indicating a strong interconnection between social media channels and donation pathways within the social programmer ecosystem [42]. Our research demonstrates that actively engaging on social media platforms to promote sponsorship opportunities for open-source development can yield fruitful outcomes. This suggests that open-source developers stand to benefit from expanding their presence and networking efforts beyond the GitHub platform. Furthermore, our study highlights the notion that publicity and visibility in the realm of open-source sponsorship need not be limited to a unidirectional flow. Rather, sponsors themselves have the potential to enhance the exposure and reach of open-source projects by publicizing their donations. In doing so, they serve as exemplars, setting a positive precedent and inspiring others to follow suit. By emphasizing these key findings, we provide compelling evidence to support the notion that using social media channels, diversifying online networks, and fostering mutual publicity between sponsors and developers can yield substantial advantages within the open-source community. These insights encourage open-source developers and sponsors alike to consider the broader potential of social media engagement and collaborative promotion to achieve their goals.

2 RESEARCH QUESTIONS

The main objective of our study is to understand the state of practice and the impact of GitHub Sponsors profile mentions on Twitter/X. The insights drawn from this study will not only contribute to the academic understanding but also have practical implications for developers, sponsors, and platforms like GitHub. Furthermore, our findings can shed light on the relationship between social activities and monetary contributions, ultimately serving to augment the appeal of developers. To guide our investigation, we present main questions and sub-questions, along with motivations and relevance.
RQ1: *How are GitHub Sponsors profiles discussed on Twitter/X?* The motivation behind RQ1 is to provide insights into the dynamics of GitHub Sponsors profile mentions, ultimately informing better strategies for developers seeking sponsorship.

RQ1.1 *What are the characteristics of tweets mentioning GitHub Sponsors profiles from organizational and personal accounts?* Understanding the language, account types, and programming languages in these tweets will enable developers to craft more appealing content for potential sponsors, ultimately enhancing engagement.

RQ1.2 *Who mentions GitHub Sponsors profiles on Twitter/X?* By identifying who engages with these profiles, sponsorship acquisition strategies can be tailored to target specific demographics.

RQ1.3 *What is the context of GitHub Sponsors profile mentions on Twitter/X?* Investigating the context in which profiles are mentioned will shed light on why tweets are used in sponsorship communication, potentially informing strategies for developers seeking sponsorship.

RQ1.4 *When are GitHub Sponsors profiles mentioned on Twitter/X?* Analyzing the timing of mentions can lead to the discovery of optimal moments for posting, which could help in securing sponsorship.
RQ2: *What is the impact of GitHub Sponsors profile mentions on Twitter/X?* Building on RQ1, RQ2 explores the effects of the dynamics uncovered, allowing us to measure and interpret their impact.

RQ2.1 *How are GitHub Sponsors profile mentions received on Twitter/X?* Understanding the reception will aid platforms like GitHub in providing targeted support and tools to developers, such as social media templates and guidelines.

RQ2.2 *How are GitHub Sponsors profile mentions discussed on Twitter/X?* By examining engagement metrics and replies, we will gain a deeper understanding of the conversations, ultimately helping in crafting more effective strategies for community engagement.

RQ2.3 *How do GitHub Sponsors profile mentions impact sponsorship?* Through a quasi-experimental approach, our goal is to provide quantitative evidence of the causal impact, which can guide both developers in improving their social media strategies and platforms in enhancing features that facilitate sponsorship acquisition.

By addressing these research questions, we aim to provide insights into the dynamics and consequences of GitHub Sponsors mentions on Twitter/X. This exploration contributes to theoretical understanding and practical strategies, offering value to the broader Open Source community.

3 RESEARCH METHODS

This section describes our methods for data collection and our quantitative and qualitative analyses.

3.1 Data Collection

In this study, we examine tweets containing links to GitHub Sponsors profiles. We successfully applied for Twitter/X's Academic Research Access [45], which offers a higher limit on the number of tweets that can be retrieved per month, and we analyzed tweets from May 2019, when GitHub Sponsors was launched, through April 2022, using `Twitter/X API v2` [44] in May 2022. The Twitter/X API provides a search function that allows for a set of query mechanisms against tweets. We use the "url" query to retrieve tweets that contain links with the specific substring "`github.com/sponsors/`" so that we ensure all the tweets are developer-related. We obtained 11,582 tweets that contain GitHub Sponsors profile links. Among these tweets, the majority (91%) were written in English, accounting for 10,531 tweets. We only use English tweets for the following quantitative and qualitative analyses, except **RQ1.1**.

3.2 Quantitative Analysis

To understand the characteristics of tweets mentioning GitHub Sponsors profiles (**RQ1.1**), we investigate written languages, types

of GitHub accounts, and primary programming languages of developers mentioned in the tweets. For written languages, we calculate the distribution of languages in tweets. In cases where Twitter/X cannot determine the language of a tweet (e.g., the tweet only contains hashtags, emojis, or links), `Undetermined` is used.

Since one GitHub Sponsors profile may appear in different tweets, we obtained distinct GitHub Sponsors profiles in tweets to collect the types of GitHub accounts and the primary programming languages of the corresponding developers. We obtained 3,766 distinct GitHub Sponsors profiles from the 11,582 tweets. For the types of GitHub accounts, we calculate the distribution of the types of GitHub accounts (i.e., personal or organizational) across all distinct GitHub Sponsors profiles in tweets. Since the URL of a GitHub Sponsors profile is organized as `https://github.com/sponsors/[username]`, we can retrieve the corresponding GitHub account using `username` in the GitHub GraphQL API [19].

The primary programming languages of the repositories can also be retrieved using the GitHub GraphQL API. Same to the previous work [35], we take the most common primary language of the repositories to which each developer contributed as the primary language of that developer. This is an approximation because we did not analyze whether the developer actually committed in that language. If the occurrences of each programming language per repository are the same, we consider the primary programming language to be `Undetermined`. The primary languages of developers identified in this way can be interpreted as the programming languages of the ecosystems to which the developers mainly contributed.

To attract potential sponsors, developers might be particularly active on GitHub around the time they advertise their GitHub Sponsors profile on Twitter/X. To investigate whether such correlations exist, we considered three time periods related to a "My GitHub Sponsors profile is live!" tweet, i.e., a week before posting this tweet, the week in which the tweet was posted, and a week after posting this tweet (**RQ1.4**). For example, if a tweet has been posted on 15 June 2020, these three periods will be from `2020-06-05` to `2020-06-11`, `2020-06-12` to `2020-06-18`, and `2020-06-19` to `2020-06-25`, respectively. Following the approach of related work [10], which used a time frame of one week before and after, our decision to adopt a one-week duration allows us to quickly assess immediate changes in productivity, engagement, and quality. This analysis involves scrutinizing short-term developer activities before and after sponsorship, facilitating a timely evaluation. We obtained 810 distinct GitHub Sponsors profiles that were posted using that template from our data set. Then we investigate different categories of contribution activities in each period, as shown below. To collect these contribution activities in a week, we retrieve them from the profile pages of the GitHub accounts as `https://github.com/[username]?tab=overview& from=[time period]&to=[time period]`.

- **opening pull request**: The profile page indicates that the GitHub account has opened pull requests, including substrings such as "*Created a pull request*", "*Opened 1 other pull request*", "*Opened [number] pull requests*", and "*Opened their first pull request*".
- **submitting pull request review**: The profile page indicates that the GitHub account has reviewed pull requests, including a substring such as "*Reviewed [number] pull requests*".

- **opening issue**: The profile page indicates that the GitHub account has opened issues, including substrings such as "*Created an issue*", "*Opened [number] other issues*", "*Opened their first issue*", and "*Opened [number] issues*".
- **opening discussion**: The profile page indicates that the GitHub account started a GitHub Discussion, including a substring such as "*Started [number] discussions*".
- **answering discussion**: The profile page indicates that the GitHub account answered a GitHub Discussion, including a substring such as "*Answered [number] discussions*".
- **committing**: The profile page indicates that the GitHub account has authored commits, including a substring such as "*Created [number] commits*".
- **contributing in private repository**: The profile page indicates that the GitHub account contributed to private repositories, including a substring such as "*[number] contributions in private repositories*".
- **creating repository**: The profile page indicates that the GitHub account created private repositories, including substrings such as "*Created [number] other repositories*", "*Created [number] repositories*", and "*Created their first repository*".
- **joining organization**: The profile page indicates that the GitHub account joined an organization, including a substring such as "*Joined the [name] organization*".

We conduct Mann-Whitney U tests to compare activities in these three periods, i.e., between a week before posting the tweet and the week when the tweet was posted, and between the week when the tweet was posted and a week after posting this tweet. To estimate the effect size of significant differences, we use Cliff's delta with the following thresholds [34]: negligible for $0 \leq |delta| < 0.147$, small for $0.147 \leq |delta| < 0.33$, medium for $0.33 \leq |delta| < 0.474$, and large otherwise.

To investigate the reception of tweets mentioning GitHub Sponsors profiles (**RQ2.1**), we analyze the popularity of tweets that mentioned GitHub Sponsors profiles on Twitter/X (number of likes, number of retweets, and number of replies). Then, we compare these interactions to tweets that contain links to other donation and crowd-funding platforms that are often used to obtain financial support for OSS development [28]: PayPal, Open Collective, and Patreon. To ensure that the tweets obtained are related to OSS, we collect tweets that contain links to at least one of these three platforms and GitHub (i.e., "`github.com`", except links to GitHub Sponsors) using `Twitter/X API v2` in the same time period for which we collected GitHub Sponsors profile tweets. We consider a link to point to a PayPal profile when it contains "`paypal.com/paypalme/`", Open Collective when it contains "`opencollective.com/`", and Patreon when it contains "`patreon.com/`", except Patreon posts (i.e., "`patreon.com/posts/`"). We exclude tweets that contain links to "`github.com/sponsors/`" and at least one of these three platforms from the 10,531 English tweets obtained. In the end, we obtained 10,440 tweets for GitHub Sponsors, four tweets for PayPal, 88 tweets for Open Collective, and 228 tweets for Patreon. Since only four tweets contain links to PayPal, we focus on comparisons between GitHub Sponsors, Open Collective, and Patreon.

We also conduct Mann-Whitney U tests to compare Twitter/X interactions between GitHub Sponsors and Open Collective, and

between GitHub Sponsors and Patreon. We use Cliff's delta with the same thresholds to estimate the effect size of significant differences.

3.3 Qualitative Analysis

For our qualitative analyses, we randomly selected a statistically representative number of tweets with a confidence level of 95% and a confidence interval of 5 to obtain 371 tweets from the initial population of 10,531 English-language tweets.

Unsurprisingly, our initial analysis revealed differences between the dynamics around tweets from users looking for sponsorship and those from users who made donations. Therefore, we categorized the 371 tweets into three different behavioral groups based on the purpose of the tweet, so we could see what kind of tweets a user would make based on their behavior:

- **looking for sponsors**: This tweet is posted by developers who publicized their profiles to look for sponsors, e.g.,"*My GitHub Sponsors profile is live! You can sponsor me to support my open source work* ❤️".
- **sponsors**: This tweet is posted by developers who sponsored others, e.g., "❤️*I'm sponsoring [username] because...*".
- **no purpose**: This tweet does not have sufficient information to decide, e.g., "*You guys make magic.*".

In the end, we identified 183 tweets from developers that were looking for sponsors, 168 tweets from sponsors, and 20 tweets from no purpose. These 351 tweets (20 tweets from no purpose are excluded) are used for answering RQ1.2–RQ1.4, and RQ2.2.

Four of the authors collaboratively took an initial look at a randomly selected subset of 30 tweets from the sample of 351 tweets, discussed which themes were present in the data and how these themes related to the research questions, and then formalized this discussion into coding schemata. For each aspect that entailed manual coding, a total of 30 tweets were independently labeled by four annotators, resulting in Cohen's kappa exceeding 0.6 for all parts and even reaching 0.94 for **RQ2.2**.

Encouraged by the initial kappa agreements, the first two authors independently coded the remaining sample of 321 tweets. Then, they recalculated kappa agreements to assess the improvement in understanding of the coding schemata after labeling the first 30 tweets. Finally, four authors engaged in collaborative discussions to attain a consensus in cases of disagreement. We attribute this stability to the fact that we had an initial discussion about all data, that tweets are relatively short, and that this particular team of authors has experience working together on qualitative data analysis from previous research projects. We describe the coding schemata related to each research question in the following paragraphs.

To investigate who mentions GitHub Sponsors profiles on Twitter/X (**RQ1.2**), we analyze the relationship between the authors of the tweets and the GitHub accounts that are linked in the tweets. Furthermore, the rationale behind having a "user" category in the aforementioned code is rooted in the goal of acquiring insights into what extent users benefit from a developer's project and are willing to voluntarily advertise the developer, thereby enabling the developer to obtain more sponsorship. The prevalence of the code "non-specific" in the results indicates that some users advertise for others without a specific purpose. Since the names of accounts on Twitter/X and GitHub do not necessarily have to be the same, we

employed qualitative analysis for this investigation. It is important to highlight that our decision not to employ automated techniques for verifying the association between Twitter/X and GitHub accounts was motivated by the realization that these techniques often fail to account for certain scenarios. For example, when a GitHub account is classified as an organization type and one of its members posts a tweet, it should be regarded as emanating from the same user. Consequently, we opt for a cautious approach, acknowledging the limitations of automated techniques and acknowledging the need for context-sensitive judgment in determining the correspondence between Twitter/X and GitHub accounts. The four annotators independently labeled 30 tweets. Then, we calculate the kappa agreement of our coding schemata from four annotators. The initial Cohen's kappa for this qualitative analysis is 0.75, which indicates 'substantial' agreement [46]. For the remaining sample of 321 tweets, Cohen's kappa is 0.78, which also indicates 'substantial' agreement, from the first two authors. Examples of the following codes are covered in our replication package, aiming to facilitate the reader's comprehension of this taxonomy.

- **same**: The author of this tweet is the same as the GitHub account that is shown on the GitHub Sponsors profile linked in the tweet, or the content of the tweet implies that they are the same developer or the author belongs to the GitHub organizational account on that GitHub Sponsors profile.
- **user**: The tweet explicitly indicates that the author of this tweet is a user of an open-source project that belongs to the GitHub account on the GitHub Sponsors profile.
- **non-specific**: There is not sufficient information to determine the relationship between the tweet author and the GitHub account.

To understand the context of tweets mentioning GitHub Sponsors profiles (**RQ1.3**), we analyze why GitHub Sponsors profiles were mentioned in tweets. Additionally, the reason to distinguish between the "advertisement with new information" and "advertisement with new functionality" categories in the aforementioned coding schemata is to enable a more nuanced analysis: the former encompasses a range of updates, including changes to users' profile descriptions and tier information whereas the latter is related to functionality in the projects they are dedicated to. The four annotators independently coded 30 tweets, achieving the initial Cohen's kappa of 0.66 or 'substantial' agreement [46]. The first two authors then independently labeled the remaining sample of 321 tweets, finally reaching Cohen's kappa of 0.84 or 'almost perfect' agreement. The following list shows the coding schema that emerged from the data. Examples of the following codes are described in detail in our replication package to help the reader understand this taxonomy.

- **generic advertisement**: This tweet advertises the tweet author's own GitHub Sponsors profile (use this code if the tweet does not fit the other advertisement categories).
- **donation appreciation**: This tweet explicitly expresses appreciation of a donation.
- **sponsor template**: This tweet contains GitHub's template for advertising one's own GitHub Sponsors profile: "*My GitHub Sponsors profile is live! You can sponsor me to support my open source work*❤️" with no or minor changes.

- **advertisement of developer**: This tweet advertises the GitHub Sponsors profile of another personal GitHub account.
- **advertisement with new functionality**: This tweet explicitly advertises the author's own GitHub Sponsors profile while mentioning new functionality of an open-source project.
- **advertisement with new information**: This tweet explicitly advertises the author's own GitHub Sponsors profile with an update.
- **sustainability**: This tweet explicitly indicates an appreciation or need for a donation for the sustainability of an open-source project, often associated with terms such as "sustainable".
- **advertisement with early access**: This tweet explicitly advertises the author's own GitHub Sponsors profile with early access to features (usually accompanied by a phrase such as "early access" and "insider")
- **income**: This tweet explicitly indicates the need for income to support one's daily life.
- **advertisement with event**: This tweet explicitly advertises a GitHub Sponsors profile with an event.
- **set example / peer pressure**: This tweet explicitly motivates others in either a positive or negative way.
- **advertisement of organization**: This tweet advertises the GitHub Sponsors profile of another organizational GitHub account.
- **donation to developer announcement**: This tweet explicitly indicates that the author of this tweet donated to the personal GitHub account in the GitHub Sponsors profile.
- **donation to organization announcement**: This tweet explicitly indicates that the author of this tweet donated to the organizational GitHub account in the GitHub Sponsors profile.
- **donation to developer template**: This tweet contains GitHub's template that indicates donation to a personal GitHub account: "💖 *I'm sponsoring [username] because...*" with no or minor changes.
- **donation to organization template**: This tweet contains GitHub's template that indicates donation to an organizational GitHub account: "💖 *I'm sponsoring [username] because...*" with no or minor changes.

To study when tweets related to GitHub Sponsors profiles occur in relation to other activities on GitHub (**RQ1.4**), we analyze the timing of such tweets. The four annotators independently coded 30 tweets, achieving the initial Cohen's kappa of 0.62 or 'substantial' agreement [46]. For the remaining 321 tweets, the first two authors reached Cohen's kappa of 0.85 or 'almost perfect' agreement. The following list shows the coding schemata that emerged from the data.

- **start**: This tweet was posted when the GitHub Sponsors profile is activated (usually accompanied by a phrase like "profile is live").
- **no specific timing**: This tweet was posted with no particular timing.
- **donation**: This tweet was posted when the author of the tweet received a donation.

- **update**: This tweet was posted when there was an update to a GitHub project or GitHub Sponsors profile.
- **reach goal**: This tweet was posted in relation to reaching a goal.
- **release**: This tweet explicitly indicates that a release of the software project has been delivered.
- **event**: This tweet was posted when an event has been announced.
- **resignation / paycut**: This tweet was posted during a change in the author's work professional situation.
- **benefit**: This tweet explicitly mentions a particular benefit.
- **activity spike**: This tweet was posted to indicate the GitHub developer was particularly active and explicitly mentions the activity spike.

To investigate the responses to tweets mentioning GitHub Sponsors profiles (**RQ2.2**), we analyze the replies to tweets mentioning GitHub Sponsors profiles. Four annotators independently annotated 30 tweets. The initial kappa agreement is 0.94, interpreted as 'almost perfect' agreement [46]. The first two authors independently annotated the remaining 321 tweets, reaching the same kappa agreement. Our coding schemata emerged from the data and is as follows. Note that examples for these codes are described in detail in our replication package to help the reader understand this taxonomy.

- **support**: The response to this tweet demonstrates endorsement or encouragement for the author, often extending beyond appreciation and indicating a willingness to assist or advocate for the author's cause.
- **appreciation of work**: The respondent acknowledges and values the author's open-source contributions and their impact, without necessarily conveying explicit support or a commitment to assist in further efforts.
- **appreciation of donation**: The respondent to this tweet appreciates the donation.
- **emoji only**: The response to this tweet only contains emoji.
- **other**: The response to this tweet does not fit into the categories above, or there is no response to this tweet.

3.4 Causal Inference

We conduct a quasi-experiment to estimate the causal impact of GitHub Sponsors profile mentions in tweets on the number of sponsors acquired (**RQ2.3**). Unlike prior studies that have conducted quasi-experiments for causal inference in software engineering by employing difference-in-differences [15, 26, 47] or CausalImpact [25], we are unable to employ these methods. This is because these methods require the values of the outcome variables in the periods before and after the treatment, but data on the number of sponsors at a given point in time were not available at the time we conducted our analysis.[1] Therefore, in this analysis, we apply a statistical matching method called propensity score matching (PSM), which attempts to estimate the effect of treatment by constructing a control group by matching each treated unit with a non-treated unit with similar characteristics [22]. PSM predicts the probability of belonging to

---

[1]We contacted the GitHub team in a public forum and they responded that they would consider making the sponsor count data publicly available; we do not provide a link to that form because of the double-anonymous submission.

the treatment and control groups based on observed predictors. Some of the studies mentioned above used PSM to prepare data for the treatment and control groups [15, 26].

To collect developers as potential members of a control group, we contacted the authors of previous work [35] to obtain the list of GitHub users who had participated in GitHub Sponsors. From 3,697 sponsored and 5,666 non-sponsored developers collected in July 2021 for the previous study, we identified 1,930 and 4,913 developers who had not deleted their GitHub accounts and whose GitHub Sponsors profiles do not appear in our tweet data (neither in their own tweets nor in tweets from others). Potential members of the treatment group are developers whose GitHub Sponsors profiles appear in the "sponsor template" tweets, that is, "My GitHub Sponsors profile is live!". By targeting only "sponsor templates", the influence of wording differences can be eliminated. We observed that "sponsor template" appears most often after "sponsor template", which are free-text tweets (see Section 4.1.3). To limit developers to those who started using GitHub Sponsors at the same period as control group developers, we collected only those developers whose GitHub Sponsors profiles appeared in such tweets by July 2021 and identified 568 developers.

The following are variables of developers used in the logistic regression to estimate the propensity score for PSM.

These variables have been used in previous related studies: for example, sponsored developers sponsor more than non-sponsored developers [35], sponsored developers form language-specific clusters that sponsor each other [35], and the number of followers is the most important feature for predicting long-term contributors [3]. All values were measured in August 2022.

- **repositories**: The number of public repositories created.
- **sponsoring**: The number of developers sponsoring.
- **openedPRs**: The number of opened pull requests.
- **reviewedPRs**: The number of reviewed pull requests.
- **followers**: The number of followers.
- **organizations**: The number of joined organizations.
- **language**: Categorical variable for the primary programming language determined by the method described in Section 3.2. The values are the top 10 languages (JavaScript, Python, PHP, C#, Go, Java, TypeScript, C++, Ruby, and C) and others (including undetermined) seen in Table 2. In regression model building, dummy variables are prepared that take a value of 0 or 1 indicating the absence or presence of a particular language.

We obtained 1,094 matched developers out of 7,411 (1, 930 + 4, 913 + 568) developers from the PSM. Figure 1 shows how the absolute mean differences have decreased as a result of the matching, from unadjusted to adjusted (unadjusted indicates all developers before matching, and adjusted indicates matched developers). None of the absolute mean differences of adjusted exceeds 0.10, which means that we obtained developer matches for the treatment and control groups with a balanced distribution of covariates [22]. This balance is a measure of the quality of the propensity score matching and we achieved the well-established and well-cited threshold [2, 27].

To estimate the impact of GitHub Sponsors profile mentions in tweets, a linear regression is performed using the above variables
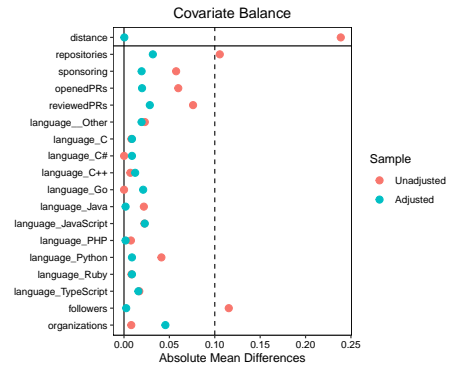

Figure 1: Covariate balance before (unadjusted) and after (adjusted) propensity score matching.

Table 1: Frequency of written languages of tweets that contain links to GitHub Sponsors profiles.

written languages	Person	Organization
English	3,074 (94%)	479 (97%)
Japanese	151 (5%)	5 (1%)
Undetermined	18 (0%)	1 (0%)
Spanish	8 (0%)	1 (0%)
Other	19 (0%)	4 (0%)
sum	3,270 (100%)	496 (100%)

and the variable treatment, which takes a value of 0 or 1 that indicates the presence or absence of general template tweets. The outcome variable is the number of sponsors obtained by each developer, measured in August 2022. Therefore, this analysis estimates the impact of tweeting "My GitHub Sponsors profile is live!" on the number of sponsors as of August 2022, for early adopters starting GitHub Sponsors and tweeting from May 2019 (GitHub Sponsors launched) through July 2021.

4 RESULTS

This section presents answers to our research questions.

4.1 RQ1: *How are GitHub Sponsors profiles discussed on Twitter/X?*

The results of the analysis of the characteristics, participants, context, and timing of tweets mentioning GitHub Sponsors profiles are presented in this section.

*4.1.1 **RQ1.1: What are the characteristics of tweets mentioning GitHub Sponsors profiles from organizational and personal accounts?*** We investigated the written languages, GitHub account types, and the primary programming languages of the developers mentioned in the tweets. These elements were categorized based on whether they originated from personal or organizational accounts. This initial analysis serves as a foundation for our subsequent in-depth investigation, offering an initial understanding of the nature of these tweets.

Written languages. Table 1 presents the frequency of written languages in tweets that contain links to GitHub Sponsors profiles.

Table 2: Frequency of GitHub account types and primary programming languages of distinct GitHub Sponsors profiles.

programming languages	Person	Organization
JavaScript	816 (25%)	89 (18%)
Python	333 (10%)	47 (9%)
PHP	309 (9%)	44 (9%)
C#	228 (7%)	22 (4%)
Go	180 (6%)	21 (4%)
Java	153 (5%)	21 (4%)
Other	1251 (38%)	252 (51%)
sum	3,270 (100%)	496 (100%)

Table 3: Frequency of relationships between tweet authors and linked GitHub Sponsors profiles.

	looking for sponsors	sponsors
same	169 (48%)	-
user	-	55 (16%)
non-specific	14 (4%)	113 (32%)
sum	183 (52%)	168 (48%)

Table 4: Frequency of context of the GitHub Sponsors profile mentions on Twitter/X.

	looking for sponsors	sponsors
generic advertisement	72 (21%)	-
donation appreciation	34 (10%)	-
sponsor template	33 (9%)	-
advertisement of developer	9 (3%)	7 (2%)
advertisement with new functionality	9 (3%)	-
advertisement with new information	8 (2%)	-
sustainability	5 (1%)	-
advertisement with early access	5 (1%)	-
income	2 (1%)	-
advertisement with event	3 (1%)	-
set example / peer pressure	2 (1%)	4 (1%)
advertisement of organization	1 (0%)	2 (1%)
donation to developer announcement	-	101 (29%)
donation to organization announcement	-	33 (9%)
donation to developer template	-	16 (5%)
donation to organization template	-	5 (1%)
sum	183 (52%)	168 (48%)

Compared to the ranks and portions of the written languages of general tweets [21], in English GitHub Sponsors profile tweets, both personal and organizational accounts make up a significantly larger portion than general English tweets (51%). Japanese tweets rank second in general tweets, comprising 5% of personal accounts and 1% of organizational accounts. Spanish also stands out as a top contributor among the top ten languages used frequently in general tweets.

GitHub account types. As seen in Table 1 and Table 2, most GitHub Sponsors profiles mentioned in tweets are associated with personal accounts, accounting for 87% of 3,766 distinct GitHub Sponsors profiles. Approximately a fifth of GitHub Sponsors profiles in tweets are associated with organizational accounts, representing 13% of distinct GitHub Sponsors profiles in the obtained tweets. According to GitHub's advanced search engine [18] in August 2022, 18,129 personal GitHub accounts had activated GitHub Sponsors, accounting for 91%. Furthermore, only 9% of all GitHub accounts (1,889) that activated GitHub Sponsors are organizational accounts. Comparing GitHub Sponsors profiles that were posted on Twitter/X and all GitHub Sponsors profiles on GitHub, they tend to share a similar trend for GitHub account types.

Programming languages. Among the 3,766 distinct GitHub Sponsors profiles mentioned in the collected tweets, JavaScript stands out as the most prominent language, with 25% of personal accounts, suggesting its popularity among individual users. Conversely, its relatively lower representation in organizational accounts (18%) may indicate a preference for other languages in professional settings. Python, with 10% of usage among personal accounts, appears to be a language of choice for individual developers, potentially due to its versatility and readability. The prevalence of Python and PHP, both of them at 9%, among organizational accounts hints at their significance in enterprise-level development projects, as seen in Table 2. In the "other" category of coding repositories, where many instances are labeled as "None", there are organizations like PJSoftCo.[2] They are a prime example of how GitHub organizations are using sponsorship funds to invest in their organization-wide documentation. Comparing these results with previous work [35], we find that, except for `Undetermined`, the top four programming languages are exactly the same. The top ten primary programming languages are the same on individual GitHub Sponsors and GitHub Sponsors profiles that were posted on Twitter/X.

---

[2]https://github.com/PJSoftCo

4.1.2 ***RQ1.2: Who mentions GitHub Sponsors profiles on Twitter/X?*** Table 3 shows the results of the coding for **RQ1.2**. As mentioned in Section 3.3, we separated tweets by purpose, distinguishing developers who mention their GitHub Sponsors profiles to look for sponsors from those who are sponsors. Developers that were looking for sponsors mentioning their own GitHub Sponsors profiles in tweets is the most frequently occurring case, accounting for 48% of the sample. However, it is also common that sponsors mention GitHub Sponsors profiles of other GitHub accounts, accounting for 32% of the sample. We observe that sponsors also explicitly mentioned GitHub Sponsors profiles of others due to dependencies or other benefits, accounting for 16% of the sample. In previous work [35], Shimada et al. showed that developers sponsoring others via GitHub Sponsors due to dependencies is the most frequent reason for sponsoring. In the context of Twitter/X, our result partially agrees with their observations.

4.1.3 ***RQ1.3: What is the context of GitHub Sponsors profile mentions on Twitter/X?*** In Table 4, the frequency of various GitHub Sponsors profile mentions on Twitter/X is presented. Most sponsors mentioned GitHub Sponsors profiles on Twitter/X in the context of donating to personal or organizational GitHub accounts, accounting for 29% and 9%, respectively. Sponsors also mentioned their donation to personal or organizational GitHub accounts using GitHub's tweet templates, accounting for 5% and 1%, respectively.

In addition to donations, developers looking for sponsors mentioned GitHub Sponsors profiles on Twitter/X to advertise their own profile (21%) or to advertise profiles of other personal GitHub accounts (3%). Specifically, developers looking for sponsors advertise

Table 5: Frequency of timing of tweets that contain links to GitHub Sponsors profiles.

	looking for sponsors	sponsors
start	62 (18%)	-
no specific timing	46 (13%)	111 (32%)
donation	35 (10%)	-
update	20 (6%)	-
reach goal	8 (2%)	-
release	6 (2%)	-
event	3 (1%)	-
resignation / paycut	3 (1%)	-
benefit	-	55 (16%)
activity spike	-	2 (1%)
sum	182 (52%)	169 (48%)

their own profile with updates on the functionality of the project (3%), updates on the profile (2%), early access features (1%), and events (1%). As with the donation, developers looking for sponsors advertised their own profiles using GitHub's tweet templates, accounting for 9%. Furthermore, a few developers looking for sponsors also posted tweets to encourage others to donate by setting an example or applying peer pressure, accounting for 1%.

Some developers looking for sponsors use Twitter/X as a channel to express appreciation to sponsors (10%). Furthermore, a few tweets from developers looking for sponsors mention GitHub Sponsors profiles in the context of sustainability of the project or the financial income of developers. In particular, we see that several tweets are posted to share GitHub Sponsors updates in the context of the Log4j vulnerability [12] that was exploited in December 2021. For example, "*It's nice to see that a month after the Log4Shell vulnerability Log4j's maintainer has 101 GitHub Sponsors instead of 3, including corporate accounts such as Amazon Web Services*".

4.1.4 ***RQ1.4: When are GitHub Sponsors profiles mentioned on Twitter/X?*** Table 5 presents the frequency of different types of timing when different types of developers mentioned GitHub Sponsors profiles on Twitter/X. Regardless of the different types of developers, we find that most of the tweets (45%) do not specify an explicit explanation of the reason for the tweet's posting at that particular time. However, we can see that some tweets (16%) were posted at a time when sponsors benefited from a project. Developers that were looking for sponsors posted those tweets during the initiation of GitHub Sponsors profiles (18%), at the time of donation (10%), or when updating projects or profiles (6%). Furthermore, some developers that were looking for sponsors posted tweets with GitHub Sponsors profile mentions when they need financial resources due to changes in their work arrangements.

As seen in Table 5, we found an interesting type of timing with regard to when GitHub Sponsors profiles were mentioned in tweets: activity spikes, i.e., a sponsor donated due to an activity spike of a developer. Inspired by this code, we conducted a quantitative study to analyze the correlations between contributions of developers and GitHub Sponsors profile mentions on Twitter/X. Table 6 presents comparisons among three periods across a set of GitHub contribution types. Since the GitHub organization account lacks information on activity, we excluded tweets that contain GitHub Sponsors from organizations. Then, we focus on tweets from distinct developers that tweeted with "My GitHub Sponsors profile

is live!". Only 810 tweets were included out of the 10,531 English tweets. The rationale behind this choice was to specifically analyze the initial reactions and sentiments expressed by users who had just enabled their GitHub Sponsors account. Our primary goal was to capture the immediate activity of individuals in this specific context. Our sample is representative of GitHub Sponsors users' initial tweets about their GitHub Sponsors account, but not of all tweets in our dataset. We observe that most of the mean values for the week in which a tweet was posted are higher than the corresponding values in the week before or after. For the contribution activities of Opening discussion, Committing, and Creating repository, there are significant differences between a week before posting this tweet and the week when the tweet was posted, and between the week when the tweet was posted and a week after posting this tweet, with at least negligible effect sizes (Committing shows small effect size). In addition to these activities, comparing a week before posting the tweet and the week when the tweet was posted, we find that developers proposed significantly more pull requests, with negligible effect size. These results indicate that when developers posted their GitHub Sponsors profile on Twitter/X, they generally contribute more actively to OSS projects.

Summary: Of the GitHub Sponsors profiles mentioned in the tweets, 87% belong to individual developers, whose top primary languages were JavaScript, Python, and PHP. Such tweets were posted by the owners of the profiles or by others who depended on the work of the developer they sponsored. Developers looking for sponsors were more active on GitHub during the week in which tweets linking to their GitHub Sponsors profile were posted.

Table 6: Comparisons among three periods of GitHub contributions

	Before			During			After		
	mean	Q3	effect size	mean	Q3	effect size	mean	Q3	effect size
Opening PR	1.12	1	**0.0929*****	**1.75**	**2**	-	1.5	1	-
Subm. PR review	1.2	0	-	**1.28**	**0**	-	1.19	0	-
Opening issue	0.61	0	-	**1.01**	**1**	-	0.84	1	-
Opening disc.	0.03	0	**0.0419*****	**0.12**	**0**	-	0.06	0	**0.0274****
Answering disc.	0.02	0	-	**0.03**	**0**	-	**0.04**	0	-
Committing	13.49	15	**0.181*****	**18.31**	**21**	-	14.4	16.75	**0.153*****
Contr. to priv. repo.	7.69	6	-	**8.39**	**7**	-	7.42	5.75	-
Creating repo.	0.52	0	**0.0938*****	**0.71**	**1**	-	0.5	0	**0.103*****
Joining org.	0.01	0	-	**0.01**	**0**	-	0.01	0	-

* p-value < 0.05; ** p-value < 0.01; and *** p-value < 0.001. The Cliff's delta effect size with thresholds [34] are highlighted in Negligible Small Medium Large The hyphen (-) is used as a placeholder for cases where p-value ≥ 0.05 indicates there is no significant difference in the comparison or when comparing to itself.

4.2 RQ2: *What is the impact of GitHub Sponsors profile mentions on Twitter/X?*

4.2.1 ***RQ2.1: How are GitHub Sponsors profile mentions received on Twitter/X?*** Table 8 presents comparisons between donation and crowd-funding platforms in Twitter interactions. For median values, Patreon tweets received the highest number of likes,

Table 7: Frequency of response to tweets.

	looking for sponsors	sponsors
other	165 (47%)	138 (39%)
appreciation of work	9 (3%)	-
support	7 (2%)	3 (1%)
appreciation of donation	2 (1%)	25 (7%)
emoji only	-	2 (1%)
sum	183 (52%)	168 (48%)

accounting for six. Additionally, Patreon tweets were retweeted twice, which is the highest number of retweets in terms of median values. The median values of replies for the three platforms are zero. In terms of likes, the p-value of Patreon vs. GitHub Sponsors is less than 0.05 and Cliff's delta is 0.265, indicating Patreon and GitHub Sponsors have a significant difference with a small effect size. In terms of retweets, Open Collective and GitHub Sponsors have a significant difference (i.e., p-value < 0.05) with a small effect size (i.e., $0.147 \leq |delta| < 0.33$). Comparing Patreon and GitHub Sponsors in terms of retweets, there is a significant difference with a small effect size. In terms of replies, we find that Patreon and GitHub Sponsors have a significant difference with a negligible effect size (i.e., Cliff's delta is 0.132). In conclusion, while the number of tweets containing GitHub Sponsors profile mentions is much larger, tweets that link to Patreon or Open Collective in the context of OSS receive more likes and retweets.

4.2.2 RQ2.2: How are GitHub Sponsors profile mentions discussed on Twitter/X?
Table 9 shows the results of our coding of replies to tweets that mentioned GitHub Sponsors profiles. We can see that most tweets (86%) do not receive a response on Twitter/X. For the remaining 14%, the majority consists of expressions of appreciation for donations (8%). Since Twitter/X is an informal communication channel, we observe that some responses consist only of one or more emoji.

4.2.3 RQ2.3: How do GitHub Sponsors profile mentions impact sponsorship?
Table 10 summarizes the regression result. As seen in the coefficient estimate of `treatment`, there is a statistically significant positive effect of GitHub Sponsors profile mentions in tweets on the number of sponsors acquired. As the average of the expected causal effect of treatment on individuals in the treatment group, called Average Treatment Effects on the Treated (ATT), we find that tweet mentions have an impact of increasing the number of sponsors by 1.22. However, note that the medians, Q3, and means for the matched treatment and control groups are {0, 2.00, 2.56} and {0, 1.00, 1.30}, respectively, indicating a skewness in the developers who obtained sponsorship, that is, the effects are not uniform.

Summary: GitHub Sponsors profile mentions have a positive impact on the number of sponsors acquired, increasing the number of sponsors by 1.22. On Twitter/X, tweets mentioning GitHub Sponsors receive fewer interactions than those mentioning Patreon or Open Collective, and most tweets do not attract replies.

5 THREATS TO VALIDITY

Subjective nature of coding. We conducted qualitative analyses of a statistically representative sample of tweets. The codes we assigned to different tweets may be inadequate due to the subjective nature of understanding the various coding schemata. To migrate this threat, we require kappa agreements of at least "substantial agreement" to ensure a common understanding of the coding schemata among all four annotators. Then, we initiated another round of coding between the first two authors for the remaining sample. By recalculating kappa agreements, we can see the improvement in understanding the various coding schemata. For example, Cohen's kappa increased from 0.62 for the first 30 tweets to 0.85 for the reaming 321 tweets in the coding of the timing of tweets. The final results are based on the codes on which the authors, after discussion, reached a consensus and collectively agreed.

Limitations in causal inference result. We compared developers with and without tweets who started GitHub Sponsors in the same period and engaged in similar activities, but we may have missed important developer characteristics other than the metrics we measured. The result is best interpreted as an increase in sponsors acquired through social activities on Twitter/X, rather than simply tweeting "My GitHub Sponsors profile is live!". In this study, we only analyzed the impact of tweets using such a template, so the impact of free-text tweets is unknown. In addition, since this analysis was conducted on early adopters, it is not possible to generalize whether similar effects will be seen in the future, so a continued analysis is needed.

Multiple GitHub Sponsors profiles in the same tweet. There is a small number of cases where the same tweet contains multiple GitHub Sponsors profile links to different GitHub accounts. Since these cases are rare (i.e., only five tweets) and to avoid confusion in our analyses, we exclude these tweets from our analyses.

Only tweets with GitHub Sponsors profiles links. Simple keyword searches would have introduced too much noise to our large-scale analysis. To avoid false positives, we only recovered tweets with GitHub Sponsors profile links, but we acknowledge that other relevant tweets without links may have been omitted.

Primary programming languages of the developers. We considered the most common primary language of the repositories to which each developer contributed as the primary language of that developer. This means it could happen for some users, for example, that the most common language of a developer's contributed repositories is Java, but the developer may only contribute the documentation of these Java repositories, whereas committing Python code to another project. Therefore, it is important to acknowledge this potential limitation in accurately capturing a developer's primary programming language through this methodology.

The number of tweets mentioning GitHub Sponsors has a different scale of data compared with other sponsorship platforms. We collected 10,440 tweets for GitHub Sponsors compared to other platforms: 4 for PayPal, 88 for Open Collective, and 228 for Patreon. It is important to recognize that this difference in data size may affect the robustness and generalizability of our conclusions.

External validity is concerned with our ability to generalize based on our results. In Section 4.1.4, we used a subset of 810 tweets from a pool of 10,531 English tweets. It is crucial to acknowledge that the

Table 8: Comparisons among donation and crowd-funding platforms in Twitter interactions.

	like		retweet		reply		#
	median	effect size	median	effect size	median	effect size	
Open Collective	4	-	1	**0.216*****	0	-	88
Patreon	**6**	**0.265*****	2	**0.278*****	0	**0.132*****	228
GitHub Sponsors	3	-	0	-	0	-	10,440

* p-value < 0.05; ** p-value < 0.01; and *** p-value < 0.001. The Cliff's delta effect size with thresholds [34] are highlighted in Negligible Small Medium Large. The hyphen (-) is used as a placeholder for cases where p-value ≥ 0.05 indicates there is no significant difference in the comparison or when comparing to itself (GitHub Sponsors).

Table 9: Frequency of response to tweets.

	looking for sponsors	sponsors
none	140 (40%)	130 (37%)
other	27 (8%)	9 (3%)
support	6 (2%)	2 (1%)
appreciation of work	6 (2%)	-
appreciation of donation	2 (1%)	25 (7%)
emoji only	-	2 (1%)
sum	182 (52%)	169 (48%)

Table 10: Causal inference impact of GitHub Sponsors Profile Mentions in "My GitHub Sponsors Profile is Live!" Tweets.

	estimate	std. error	p
treatment	**1.22**	0.452	**0.00681**
repositories	-0.000818	0.00209	0.696
sponsoring	1.12	0.194	1.01e-8
openedPRs	0.000432	0.000457	0.345
reviewedPRs	0.00301	0.00140	0.0325
followers	0.00271	0.000345	8.42e-15
organizations	-0.0637	0.0870	0.465
language_JavaScript	-1.45	0.657	0.0279
language_Python	-1.18	0.856	0.168
language_PHP	-0.343	0.906	0.705
language_C#	0.559	0.973	0.566
language_Go	-0.933	1.02	0.360
language_Java	-1.66	1.15	0.149
language_TypeScript	-0.692	1.04	0.505
language_C++	-1.07	1.44	0.458
language_Ruby	-1.01	1.50	0.500
language_C	-0.659	1.21	0.586

chosen subset may not fully represent the broader spectrum of reactions across all types of tweets related to GitHub Sponsors. Users who express their thoughts in different formats or use alternative phrases may not be fully captured in our analysis.

6   DISCUSSION

This section presents implications and future work from this study.

Implications. We categorized the practical implications for diverse groups of individuals by offering tailored guidance and recommendations that align with the specific concerns and interests of each stakeholder group.

- *Developers seeking sponsorship*: our study shows that mentioning GitHub Sponsors profiles in tweets has a positive impact on the number of sponsors acquired. The finding that the number of sponsors acquired increased depending on whether they tweeted, is evidence of the importance of social

media and should encourage developers to go beyond the GitHub platform in order to attract sponsorship. Additionally, our research reveals many different types of messages surrounding GitHub Sponsors in various contexts, providing insights that might assist others in crafting their own effective social media strategies for sponsorship engagement.

- *Developers interested in sponsoring*: Within our sample, approximately half of the participants are sponsors. This finding underscores the importance of encouraging users who depend on OSS projects to actively promote the developers they rely on, even if the sponsorship amount is not substantial. Engaging in social media promotion can significantly enhance the visibility of these developers, allowing their exceptional work to reach a wider audience and garner increased recognition.

- *Companies*: The relationship between companies and OSS projects is undergoing a pivotal change, largely driven by an expanding recognition of sustainability issues inherent in OSS. Instead of merely expressing dissatisfaction with the lack of sustainability in these projects, our study offers evidence that a two-pronged approach of corporate sponsorship and active social media engagement could be an effective strategy for businesses. This strategy allows them to constructively engage with OSS projects they rely upon, particularly those struggling with sustainability, thereby addressing their concerns and contributing to potential solutions.

Future Work. As our study is positioned as an early adopter study, we have not yet obtained conclusive evidence of a significant impact of financial support on OSS sustainability at this stage. Therefore, further investigation of the potential impact of financial support on sustainability is needed.

We have focused on Twitter/X as the starting point for our exploration. For future research, there is significant value in extending our analysis to encompass posts from multiple social media platforms (e.g., Facebook, Reddit) to gain a more comprehensive understanding of these dynamics.

As part of our investigation into **RQ2.1**, we found that tweets linking exclusively to GitHub Sponsors were more common, and among those platforms, GitHub is the only one that provides Twitter/X templates for developers looking for sponsors and Twitter/X templates for sponsors. However, GitHub Sponsors received fewer responses compared to tweets promoting alternative sponsorship platforms. Since GitHub Sponsors launched 4–6 years later than Open Collective and Patreon, so it had less time to solidify its position and gain widespread awareness. Further research is needed

to determine the importance of the template if such a sponsorship platform provides a template when users are trying to advertise on social media, and strategies for increasing the response and engagement of tweets containing only links to GitHub Sponsors.

In light of our findings regarding the skewness in the developers who obtained sponsorship (Section 4.2.3), factors other than tweets may play an important role in sponsor acquisition. Thus, further research is needed to explore and identify these additional factors that contribute to the sponsor acquisition process.

In subsequent studies, it would be valuable to investigate the specific strategies and practices employed by organizations when leveraging social media platforms to disseminate project updates. By examining the relationship between these practices and the resulting engagement levels within the community, we can gain insights into the effectiveness of such approaches and their potential for enhancing community involvement. Moreover, it is worth considering a more in-depth investigation into how the domain and functionalities of open-source projects can impact and guide the dynamics of sponsorship, such as evaluations of a project's sustainability and its sponsorship status [32].

7 RELATED WORK

In this section, we situate our work with respect to the literature on donations and one potential advertising channel, Social Media.

Donation. OSS development heavily relies on volunteer contributions, as highlighted in a recent GitHub survey [17], revealing that just 23% of respondents contribute to open source as part of their job description. Despite more employees being paid for contributing to OSS projects during work hours [33], developers still perceive compensation asymmetry in OSS projects [1]. OSS projects that are distributed unequally may fail if they are mismanaged and financial benefits are a factor in the sustainability of OSS projects [1]. Donation is one of the common ways to obtain these financial benefits [13] to support OSS projects, in addition to Bounty [13]. In a mixed-method empirical study, Overney et al. [28] found that only a few projects (0.04–0.2%) ask for donations, primarily using platforms like PayPal and Patreon. These projects tend to be more active, more mature, and more popular.

Recently, Zhou et al. [49] explored donations on the Open Collective platform that support open-source projects. They indicated the influence of individual donors; although corporate donors tend to donate more money than individual donors for an individual donation, the total donation amount from individual donors is greater than corporate donors. However, corporate collectives are more likely to receive a larger total donation amount than individual collectives. Regarding the study on GitHub Sponsors, Shimada et al. [35] revealed that developers typically do not have channels at their disposal to attract sponsors and communicate with those who might be interested in donating. Zhang et al. [48] discovered that sponsorship through GitHub Sponsors has a short-term impact on developers' activities. Their survey highlighted key challenges, including the difficulty of attracting sponsorship and the absence of corporate support.

Social Media. Social media channels are one way of communicating and advertising in the world of developers. Different social media channels play different roles and have different impact on OSS projects, e.g., facilitating communication [5, 38], awareness of the status of other developers [4, 6], gaining attention [8, 23, 43], and attracting new contributors [7, 23, 31].

Researchers studied the use of microblog services such as Twitter/X in software development [8]. The tweets of developers differ from those of the general public in terms of length, use of URLs, and @-mentions, and software microbloggers are more tightly knit than general microbloggers [9, 41]. Twitter/X is widely adopted in the software engineering community [5, 39]. Tian et al. [40] found that knowledge sharing, technical discussion, and software product updates are the most frequent categories of developers' tweets. Fang et al. [14] proposed an approach to cross-link users on Twitter/X and GitHub; they observed that tweeting patterns appear in tweets from different developer roles when including GitHub links in their tweets (e.g., repository owners prompt their projects instead of discussing specific software artifacts).

For the impact of Twitter/X, Singer et al. [36] indicated that Twitter/X can help developers become aware of industry changes, learn, and build work relationships in communities. Mezouar et al. [24] found that tweets from end users can lead to early discovery of bugs in web browsers. Fang et al. [15] explored the causal effects of Twitter/X on the attraction of stars and contributors by analyzing tweets that contain links to GitHub repositories. They found that Twitter/X has a statistically significant and sizeable effect to help make projects popular (i.e., stars) but only a small effect to attract new contributors (i.e., commits). Moreover, these newly attracted contributors showed to be more active in OSS projects when they had prior Twitter/X interactions with the tweet authors.

8 CONCLUSION

There are several platforms that enable contributions to open-source software developers, but attracting sponsors in order to ensure project sustainability remains a challenge. To understand the impact of Twitter/X on helping OSS developers attract sponsors, we conducted quantitative and qualitative analyses of more than 10,000 tweets containing links to GitHub Sponsors profiles. We find that such tweets have a significant positive effect on the acquisition of sponsors, and that developers contribute more OSS work than usual to attract potential sponsors during the week in which they posted tweets that link to their own GitHub Sponsors profile.

Open-source developers who maintain an active presence on social media can attract donations that help sustain their projects. Our findings suggest that social media channels and donation channels are linked in the social programmer ecosystem and will continue to grow in importance for the sustainability of open source software.

DATA AVAILABILITY

The replication package includes scripts and data set, which is available at https://doi.org/10.5281/zenodo.10461383 and https://github.com/NAIST-SE/GHSponsorsX

REFERENCES

[1] Arzoo Atiq and Arvind Tripathi. 2016. Impact of financial benefits on open source software sustainability. In *Proceedings of 37th International Conference on Information Systems (ICIS '16)*. 10.

[2] Peter C. Austin. 2009. Balance diagnostics for comparing the distribution of baseline covariates between treatment groups in propensity-score matched samples. *Statistics in Medicine* 28, 25 (2009), 3083–3107. https://doi.org/10.1002/sim.3697 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/sim.3697

[3] Lingfeng Bao, Xin Xia, David Lo, and Gail C. Murphy. 2021. A Large Scale Study of Long-Time Contributor Prediction for GitHub Projects. *IEEE Transactions on Software Engineering* 47, 6 (2021), 1277–1298. https://doi.org/10.1109/TSE.2019.2918536

[4] Andrew Begel, Robert DeLine, and Thomas Zimmermann. 2010. Social media for software engineering. In *Proceedings of the FSE/SDP workshop on Future of software engineering research*. 33–38.

[5] Sue Black, Rachel Harrison, and Mark Baldwin. 2010. A survey of social media use in software systems development. In *Proceedings of the 1st Workshop on Web 2.0 for Software Engineering*. 1–5.

[6] Sue Black and Joanne Jacobs. 2010. Using Web 2.0 to improve software quality. In *Proceedings of the 1st Workshop on Web 2.0 for Software Engineering*. 6–11.

[7] Kelly Blincoe, Jyoti Sheoran, Sean Goggins, Eva Petakovic, and Daniela Damian. 2016. Understanding the popular users: Following, affiliation influence and leadership on GitHub. *Information and Software Technology* 70 (2016), 30–39.

[8] Hudson Silva Borges and Marco Tulio Valente. 2019. How do developers promote open source projects? *Computer* 52, 8 (2019), 27–33.

[9] Gargi Bougie, Jamie Starke, Margaret-Anne Storey, and Daniel M German. 2011. Towards understanding twitter use in software engineering: preliminary findings, ongoing challenges and future questions. In *Proceedings of the 2nd international workshop on Web 2.0 for software engineering*. 31–36.

[10] Andrea Capiluppi and Daniel Izquierdo-Cortázar. 2013. Effort estimation of FLOSS projects: a study of the Linux kernel. *Empirical Software Engineering* 18 (2013), 60–88.

[11] Open Collective. 2023. Raise and spend money with full transparency. — opencollective.com. https://opencollective.com/. [Accessed 28-Jun-2023].

[12] CVE. 2021. CVE - CVE-2021-44228 — cve.mitre.org. https://cve.mitre.org/cgi-bin/cvename.cgi?name=CVE-2021-44228. [Accessed 28-Jun-2023].

[13] Nadia Eghbal. 2019. A handy guide to financial support for open source. https://github.com/nayafia/lemonade-stand. Accessed: https://opensourcesurvey.org/2017/.

[14] Hongbo Fang, Daniel Klug, Hemank Lamba, James Herbsleb, and Bogdan Vasilescu. 2020. Need for tweet: How open source developers talk about their github work on twitter. In *Proceedings of the 17th International Conference on Mining Software Repositories*. 322–326.

[15] Hongbo Fang, Hemank Lamba, James Herbsleb, and Bogdan Vasilescu. 2022. "This is Damn Slick!": Estimating the Impact of Tweets on Open Source Project Popularity and New Contributors. In *Proceedings of the 44th International Conference on Software Engineering (ICSE '22)*. 2116–2129. https://doi.org/10.1145/3510003.3510121

[16] Gitcoin. 2023. Grants — gitcoin.co. https://gitcoin.co/grants/. [Accessed 28-Jun-2023].

[17] GitHub. 2017. Open Source Survey 2017. https://opensourcesurvey.org/2017/. Accessed: 2022-08-18.

[18] GitHub. 2023. Build software better, together — github.com. https://github.com/search/. [Accessed 28-Jun-2023].

[19] GitHub. 2023. GitHub GraphQL API documentation - GitHub Docs — docs.github.com. https://docs.github.com/en/graphql. [Accessed 28-Jun-2023].

[20] Giveth. 2023. Giveth: Welcome to the Future of Giving — giveth.io. https://giveth.io. [Accessed 28-Jun-2023].

[21] Lichan Hong, Gregorio Convertino, and Ed Chi. 2011. Language matters in twitter: A large scale study. In *Proceedings of the international AAAI conference on web and social media*, Vol. 5. 518–521.

[22] Guido W. Imbens and Donald B. Rubin. 2015. *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. Cambridge University Press. https://doi.org/10.1017/CBO9781139025751

[23] Danaja Maldeniya, Ceren Budak, Lionel P Robert Jr, and Daniel M Romero. 2020. Herding a Deluge of Good Samaritans: How GitHub Projects Respond to Increased Attention. In *Proceedings of The Web Conference 2020*. 2055–2065.

[24] Mariam El Mezouar, Feng Zhang, and Ying Zou. 2018. Are tweets useful in the bug fixing process? an empirical study on firefox and chrome. *Empirical Software Engineering* 23, 3 (2018), 1704–1742.

[25] Emerson Murphy-Hill, Edward K. Smith, Caitlin Sadowski, Ciera Jaspan, Collin Winter, Matthew Jorde, Andrea Knight, Andrew Trenk, and Steve Gross. 2019. Do Developers Discover New Tools on the Toilet?. In *Proceedings of the 41st International Conference on Software Engineering* (Montreal, Quebec, Canada) *(ICSE '19)*. IEEE Press, 465–475. https://doi.org/10.1109/ICSE.2019.00059

[26] Keitaro Nakasai, Hideaki Hata, and Kenichi Matsumoto. 2019. Are Donation Badges Appealing?: A Case Study of Developer Responses to Eclipse Bug Reports.

[27] Sharon-Lise T. Normand, Mary Beth Landrum, Edward Guadagnoli, John Z. Ayanian, Thomas J. Ryan, Paul D. Cleary, and Barbara J. McNeil. 2001. Validating recommendations for coronary angiography following acute myocardial infarction in the elderly: A matched analysis using propensity scores. *Journal of Clinical Epidemiology* 54, 4 (2001), 387–398. https://doi.org/10.1016/S0895-4356(00)00321-8

[28] Cassandra Overney, Jens Meinicke, Christian Kästner, and Bogdan Vasilescu. 2020. How to Not Get Rich: An Empirical Study of Donations in Open Source. In *Proceedings of ACM/IEEE 42nd International Conference on Software Engineering (ICSE '20)*. 1209–1221.

[29] Patreon. 2023. Creativity powered by membership | Patreon — patreon.com. https://www.patreon.com/. [Accessed 28-Jun-2023].

[30] PayPal. 2023. Digital Wallets, Money Management, and More — paypal.com. https://www.paypal.com/. [Accessed 29-Jun-2023].

[31] Huilian Sophie Qiu, Yucen Lily Li, Susmita Padala, Anita Sarma, and Bogdan Vasilescu. 2019. The signals that potential contributors look for when choosing open-source projects. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 1–29.

[32] Uzma Raja and Marietta J. Tretter. 2012. Defining and Evaluating a Measure of Open Source Project Survivability. *IEEE Transactions on Software Engineering* 38, 1 (2012), 163–174. https://doi.org/10.1109/TSE.2011.39

[33] Dirk Riehle, Philipp Riemer, Carsten Kolassa, and Michael Schmidt. 2014. Paid vs. volunteer work in open source. In *Proceedings of 47th Hawaii International Conference on System Sciences (HICSS '14)*. 3286–3295.

[34] Jeanine Romano, Jeffrey D Kromrey, Jesse Coraggio, and Jeff Skowronek. 2006. Appropriate statistics for ordinal level data: Should we really be using t-test and Cohen'sd for evaluating group differences on the NSSE and other surveys. In *annual meeting of the Florida Association of Institutional Research*, Vol. 177. 34.

[35] Naomichi Shimada, Tao Xiao, Hideaki Hata, Christoph Treude, and Kenichi Matsumoto. 2022. GitHub Sponsors: Exploring a New Way to Contribute to Open Source. In *Proceedings of the 44th International Conference on Software Engineering (ICSE '22)*. 1058–1069. https://doi.org/10.1145/3510003.3510116

[36] Leif Singer, Fernando Figueira Filho, and Margaret-Anne Storey. 2014. Software engineering at the speed of light: how developers stay current using twitter. In *Proceedings of the 36th International Conference on Software Engineering*. 211–221.

[37] GitHub Sponsors. 2023. GitHub Sponsors — github.com. https://github.com/sponsors. [Accessed 28-Jun-2023].

[38] Margaret-Anne Storey, Christoph Treude, Arie van Deursen, and Li-Te Cheng. 2010. The impact of social media on software engineering practices and tools. In *Proceedings of the FSE/SDP workshop on Future of software engineering research*. 359–364.

[39] Margaret-Anne Storey, Alexey Zagalsky, Fernando Figueira Filho, Leif Singer, and Daniel M German. 2016. How social and communication channels shape and challenge a participatory culture in software development. *IEEE Transactions on Software Engineering* 43, 2 (2016), 185–204.

[40] Yuan Tian, Palakorn Achananuparp, Ibrahim Nelman Lubis, David Lo, and Ee-Peng Lim. 2012. What does software engineering community microblog about?. In *2012 9th IEEE Working Conference on Mining Software Repositories (MSR)*. IEEE, 247–250.

[41] Yuan Tian and David Lo. 2014. An exploratory study on software microblogger behaviors. In *2014 IEEE 4th Workshop on Mining Unstructured Data*. IEEE, 1–5.

[42] Christoph Treude, Fernando Figueira Filho, Brendan Cleary, and Margaret-Anne Storey. 2012. Programming in a socially networked world: the evolution of the social programmer. In *The Future of Collaborative Software Development*. 1–3.

[43] Asher Trockman, Shurui Zhou, Christian Kästner, and Bogdan Vasilescu. 2018. Adding sparkle to social coding: an empirical study of repository badges in the npm ecosystem. In *Proceedings of the 40th international conference on software engineering*. 511–522.

[44] Twitter. 2022. Twitter API Documentation — developer.twitter.com. https://developer.twitter.com/en/docs/twitter-api. [Accessed 28-Jun-2023].

[45] Twitter. 2022. Twitter API for Academic Research | Products — developer.twitter.com. https://developer.twitter.com/en/products/twitter-api/academic-research. [Accessed 28-Jun-2023].

[46] Anthony Viera and Joanne Garrett. 2005. Understanding Interobserver Agreement: The Kappa Statistic. *Family medicine* (2005).

[47] Longqi Yang, David Holtz, Sonia Jaffe, Siddharth Suri, Shilpi Sinha, Jeffrey Weston, Connor Joyce, Neha Shah, Kevin Sherman, Brent Hecht, and Jaime Teevan. 2022. The effects of remote work on collaboration among information workers. *Nature Human Behaviour* 6, 1 (01 Jan 2022), 43–54. https://doi.org/10.1038/s41562-021-01196-4

[48] Xunhui Zhang, Tao Wang, Yue Yu, Qiubing Zeng, Zhixing Li, and Huaimin Wang. 2022. Who, What, Why and How? Towards the Monetary Incentive in Crowd Collaboration: A Case Study of Github's Sponsor Mechanism. In *CHI Conference on Human Factors in Computing Systems*. 1–18.

[49] Jiayuan Zhou, Shaowei Wang, Yasutaka Kamei, Ahmed E Hassan, and Naoyasu Ubayashi. 2022. Studying donations and their expenses in open source projects: a case study of GitHub projects collecting donations through open collectives. *Empirical Software Engineering* 27, 1 (2022), 1–38.

IEEE Software 36, 03 (2019), 22–27.