The Effect of Knowledge Sharing on Open Source Contribution: A Multi-platform Perspective

Vivek Kumar Singh
University of Missouri-St. Louis
vsingh@umsl.edu

Saurav Chakraborty
University of South Florida
sauravc@mail.usf.edu

Arjun Kadian
University of South Florida
arjunkadian@mail.usf.edu

Abstract

Open source software (OSS) community plays a key role in contemporary software development. However, there is a need to better understand the factors which influence individuals' voluntary contribution on open source platforms. In this paper, we investigate how different types of knowledge sharing affect an individuals' contribution towards open source projects. We further refine knowledge sharing taxonomy by classifying explicit knowledge sharing into two sub-types – strong explicit knowledge sharing and weak explicit knowledge sharing, depending on the extent of interpersonal interaction required for knowledge transfer. In this paper, we take a multi-platform perspective – we collect data from GitHub – the biggest online platform to host open source software development, and Gitter – an open source instant messaging and chat room application designed for developers. We map the user identities across these two platforms. We analyze monthly panel data for the year 2017 consisting of 3,695 individuals. The results demonstrate that both strong and weak explicit knowledge sharing have positive relationship with open source contribution. Moreover, the tacit knowledge sharing positively moderates these relationships. Our paper extends the theoretical understanding of different knowledge sharing types and their inter-relationship, and their respective impact on contribution. Our findings have important implications for the OSS community, and especially help OSS platform designers get a better understanding of the symbiosis between different OSS platforms.

1. Introduction

Open source software (OSS) development market is expanding and is presenting an important software development model for emerging computing fields such as Big Data and Artificial Intelligence. Organizations are adopting open source model to build software in these fields. However, even after decades of prior research, understanding the factors, which motivate individuals to contribute towards open source software, is a key research question [6]. Specially, there is paucity of research in understanding how voluntary knowledge sharing affects contribution in open source community.

Knowledge sharing has always been important to information technology (IT) organizations and IT knowledge workers. For IT organizations, effective knowledge sharing can lead to improved work performance and competitive advantages [7] [11]. Despite the rich literature on knowledge sharing within organizations and open source communities [17] [21] [11] [12] [15] [23] [38], very little attention has been paid to the sharing of tacit knowledge. Furthermore, the interaction between the different knowledge sharing types and its impact on individual's contribution have not been studied.

Knowledge sharing in organizations [11] [12] and virtual communities [15] have always been of interest to information systems (IS) researchers [11] [12] [8] [9]. Open source development communities are different from traditional IT organizations in that members self–identify into them and make decisions without reference to any legally binding rules or obligations [10]. In essence, knowledge owners may participate in such communities without any obligations to engage in knowledge sharing. However, the effect of voluntary knowledge sharing on contribution has not been examined in the prior research. This research is important and helps develop a theoretical understanding of *voluntary knowledge sharing* in online communities.

There are two types of knowledge sharing presented in the literature – explicit knowledge sharing and tacit knowledge sharing [37]. Explicit knowledge sharing refers to sharing of knowledge that can be codified and written in symbolic or written form [11]. For example, sharing of knowledge through manuals and reports via knowledge management systems (KMS) is a form of explicit knowledge sharing. On the other extreme, tacit

HICSS

knowledge sharing refers to knowledge that still resides with the knowledge owner and has not yet been expressed or codified [11]. For instance, the experiences of a knowledge owner finishing a task, requiring interpersonal communication for its transfer is a form of tacit knowledge sharing.

Individuals share knowledge voluntarily in online communities. However, the knowledge sharing is limited by the functionalities of the computer mediated communication system. The individuals can transfer explicit knowledge more easily than tacit knowledge. The explicit knowledge can be easily codified and expressed in form of posts without any requirement for a real-time conversation. On the other hand, tacit knowledge needs more effort and requires real-time conversation between individuals for its transfer to take place. Traditionally, open source platforms provide tools for explicit knowledge sharing such as project Wikipedia page (also commonly known as wiki page). Recently, these platforms also provide integration with external applications such as online chat applications for tacit knowledge sharing. However, there is a paucity of research examining the effect of these large-scale integration of platforms for efficient and effective knowledge transfer, on the contribution in open source community. Therefore, we address following research question in this paper:

Research Question: How different forms of voluntary knowledge sharing types impact open source contribution?

Our study has both theoretical and practical implications. We extend the knowledge sharing taxonomy and refine our understanding of how different knowledge sharing types affects contribution in OSS. Moreover, we also examine the interaction among different knowledge sharing types. In practical implication, ours is the first study to empirically demonstrate that there is need to integrate multiple OSS platforms for different forms of knowledge sharing.

2. Literature Review and Background

Our study draws mostly upon three streams of research, namely open source software development, knowledge contribution, and knowledge sharing in online communities. Each of these streams of literature share many commonalities, our aim is to extract the dynamics of the inter-connectedness of the topics under investigation. We present our findings in the sub-sections below.

2.1. Open Source Software Development:

The emergence of open source as a platform has changed how developers perceive the technological landscape [24]. The rise of open source has led to a manifold increase in collaboration within the technology developer community [19]. Thus, the commercial aspect of open source software has been of interest to researchers, leading to OSS 2.0 phenomenon [1]. The three key components of OSS ecosystem identified are software, community, and license [2]. There has been a significant amount of research, which has examined the role of developers in organizations and their participation in external open source projects [3]. There have been studies, which investigated specific open source software [4]. Thus, understanding the motivation of open source developers is an important aspect of OSS research [5].

One of the fundamental questions pertaining to OSS development has been to identify the factors motivating developers to participate in OSS. Past researchers have employed surveys to study this dynamic. Hertel et al. [25] conducted a survey of 141 web participants to uncover three kinds of motivation: pragmatic, social and hedonic. Other studies [26] also found economic factors, such as profit motivations, hierarchical co-ordination without proprietary rights and diffusion in environments dominated by proprietary standards, to be a strong motivating factor for developers to collaborate on OSS projects.

Considering the development of software to be a creative problem, finding a solution requires significant amount of time and effort [30] [27] [28] [29]. It is often called 'collective invention' via collaboration [29]. Thus, it often leads to knowledge creation and/or augmentation. There is a definite benefit when developers collaborate and come up with better solutions. Singh [30] claim that even though open source projects are public, it is the partaking developers' relationship within the open source forum or community, which determines the project's success and application. Therefore, it can be argued that the developers are often selective about which projects they participate in and in turn, contribute to.

Although different facets of the OSS environment have been studied and analyzed, there is still a lack of consensus on what motivates developers to participate in OSS. We believe that this may be due to the strong heterogeneity of projects on these platforms. Little is known regarding the intrinsic nature of such platforms and how developers interact with each other. Our study aims to bridge this gap by analyzing the factors affecting how individual developers contribute, using two of the biggest OSS platforms in the world. The dataset thus generated, enables us to study how knowledge is created and shared using these platforms, which to the best of our knowledge has not been studied before.

2.2. Knowledge Sharing in Online Communities

Knowledge sharing can be categorized either as explicit or tacit [34]. Tacit knowledge and explicit knowledge have been studied extensively in the organizational science literature. Michael [31] first introduced the concept of tacit knowledge. Tacit knowledge is information that has a 'personal' quality which makes it difficult to document. According to Nonaka [34], it is deeply rooted in action, commitment and involvement in a specific context. Thus, there is difficulty in interpreting and transferring such knowledge from one individual to another [33]. Explicit knowledge, on the other hand, can be easily expressed in writing. It is knowledge, which is based on common understanding and thus, can be easily coded and transferred from one individual to another. Let us now have a look at how we are contextualizing these two different types of knowledge transfers for open source environment.

There has been a significant amount of research done in order to uncover how explicit knowledge can be stored or transferred to other individuals. By its definition, explicit knowledge refers to knowledge that can be transferred in a formal and systematic language [34] [35]. Therefore, any form of written document that proliferates information in a direct manner, which will be easily understandable by all, can be classified as explicit knowledge sharing. On the Wikipedia page of a GitHub project, developers provide information related to software documentation, usage manual, and other information required for the development and usage of the project. Thus, the knowledge sharing on Wikipedia page is a form of explicit knowledge. Similarly, on the GitHub platform developers can communicate by raising 'issues.' The issue section of the project lists out all the problems that developers are facing, and participants can respond to such aforementioned issues by commenting under the same. Even though this section does not follow a formal and systematic language as the Wikipedia page of the project, the participants still need to communicate using a generic set of rules while posting. Thus, it is not a form of personalized knowledge sharing. Hence, we propose that Wiki edits as a medium of strong explicit knowledge transfer, whereas the issues are a medium of weak explicit knowledge transfer in the context of open source.

Tacit knowledge is knowledge that is not explicitly explicated [36]. It is knowledge that an individual gain in the form of experience, know-how, and insights [38]. Therefore, it is difficult to convey in a formal setting, thereby distinguishing itself from explicit knowledge sharing. As such tacit knowledge transfer is much more difficult. Some researchers argue that tacit knowledge can only be transferred through up close observation, demonstration and hands-on experience [39] [37]. Others propose that transfer of tacit knowledge requires regular interaction and experience sharing between team members [32] [40]. This kind of learning mostly takes place through informal contact [41] and thus, differs from the two kinds of aforementioned interactions. Panahi [38] and Marwick [43] argue that chat rooms and discussion forums provide the opportunity for real time interactions that can facilitate tacit knowledge sharing online. In the context of open source, Gitter provides a platform where developers can interact with other developers via instant chat messages in different chatrooms. The real-time chat messaging provides the perfect opportunity for tacit knowledge sharing. As the chatrooms can be customized to link with different GitHub projects and repositories, it presents us with an ideal opportunity to investigate how these different knowledge sharing mechanisms interact with each other.

Although studies in the past have examined different forms of knowledge sharing, we posit that, in the context of open source, the synchronicity of message transfer medium determines the type of knowledge which can be transferred through the medium. To the best of our knowledge, studies in the past have not explored the synchronicity of medium for message transfer in knowledge sharing literature. Moreover, we further extend the explicit knowledge sharing taxonomy based on message synchronicity of the medium. Furthermore, based on message synchronicity and message personalization, we differentiate tacit knowledge sharing from explicit knowledge sharing in two aspects: 1) tacit knowledge sharing is highly personalized and 2) it is also synchronous in nature as it happens in real time.

2.3. Knowledge Contributions in OSS

Since the rise of open source software development platform, researchers have been trying to understand what motivates individuals to voluntarily share their expertise in online communities. OSS communities have been under the purview of multiple researchers trying to identify key drivers of knowledge contributions in online communities [14] [18] [23] Multiple social psychological perspectives have been applied to study the same. Enhanced professional reputation has been found by multiple studies to be a key driving influence on individuals helping strangers in an online platform [20] [6]. Wasko and Faraj [21] also note that altruism, generalized reciprocity and community interest to be significant motivators of contributing knowledge. Economic factors have also been found by multiple studies to have a significant impact on contribution [16]. Peddibhotla [44] find other features

like social affiliation, professional self-expression and social capital as important drivers of contribution to online forums. Like reputation enhancement, social capital was also found by multiple other studies to be a key motivator for voluntary contribution to OSS communities [15] [13].

Apart from personal perceptions of the individual themselves, many researchers have also looked at features of the platforms, which enhance participation. For example, [22] observe that size of the community is a key feature. Both studies observe that smaller communities see greater amount of user participation. The aesthetics of the OSS platform have an influence on the member participation. Design factors on the user interface like visibility and reputation have been found to be effective in amplifying user's participation.

Moqri [6] argues that even though there has been a multitude of studies in the context, there is a lack of empirical agreement. This has been attributed to two reasons: a lack of consistency regarding the measures being used and the extreme diversity that exists among OSS projects and respective platforms. Projects may vary depending upon scale, application area, and required skillset amongst various others. Our work aims to address this issue by analyzing open source contribution over multiple projects throughout 2017, on two of the biggest platforms in the world. Thus, our study aims to contribute to the literature by providing a theoretical model that evaluates how different factors affect open source contribution.

3. Theory and Hypothesis Development

As stated before, multiple studies have been conducted to identify what factors motivate developers to share their knowledge or contribute to OSS projects. Many studies have used these terms interchangeably, leading to an ambiguity regarding the relationship between these two concepts. In this study, we propose a theoretical model, which can help understand the relationship between knowledge sharing and contribution in the context of open source communities.

As knowledge sharing is a medium for developers to collaborate and augment their knowledge, we posit that more knowledge sharing will lead to greater contribution. We expect there will be an increase in the developer's contribution level if the individual gets clarifications about doubts regarding the overall direction of the project the person is participating in. Such information is provided in the Wikipedia page of the Github repository. The knowledge sharing via Wikipedia is a form of "strong explicit" knowledge sharing since it does not require any interpersonal communication. In light of above, we posit our first hypothesis.

Hypothesis 1: Strong explicit knowledge sharing will have a positive effect on developer's contribution level.

In addition, when a developer works on a project, they have specific queries that are more granular and pertain to particular module (s) of the project. These kinds of queries are addressed by raising 'Issues'- a form of forum-based discussion to receive clarification. The issues are different from Wikipedia because it requires interpersonal communication in form of a asynchronous discussion. Therefore, we conceptualize issues as a form "weak explicit" knowledge sharing in comparison to Wikipedia which is a form of strong explicit knowledge sharing. This leads us to our second hypothesis.

Hypothesis 2: Weak explicit knowledge sharing will have a positive effect on developer's contribution level.

As the issues raised may or may not get an immediate response on GitHub, it is important to have an avenue for more real-time interaction. This is provided by the Gitter platform where developers interact and collaborate in synchronously. We posit the conversations on the Gitter platform will enable a developer to make more contributions and thus, we propose our third hypothesis.

Hypothesis 3: Tacit knowledge sharing will have a positive effect on developer's contribution level.

The chat platform provided by Gitter also allows developers to mention the repository that they are currently working on. This is provided by having chat rooms based on specific repositories. For example, Scikit-learn, a large-scale open source project on GitHub for machine learning, has its own dedicated chat room in Gitter. The Gitter chat room is synchronized with GitHub to reflect any changes that have been made in Wikipedia of the GitHub repository. The Wikipedia of the repository can be referenced using the URL of the web page. Thus, we posit that the knowledge attained by the developers on the chat rooms will help them to get a better understanding of the project's overall context and goal. In light of above, we hypothesize the following.

Hypothesis 4a: Tacit knowledge sharing will positively moderate the effect of strong explicit knowledge sharing on developer's contribution level.

In addition, on the Gitter platform, within a project dedicated chat room, it is possible to chat about a specific issue by referencing it with hash key '#' key (see right panel in Figure 2). Every issue generated on GitHub forum has an associated hash number, and this number can be used to reference the said issue on Gitter. This allows developers to bring their queries in real-time to the chat room and get real-time feedback regarding the same. Hence, we arrive at the following hypothesis.

Hypothesis 4b: Tacit knowledge sharing will positively moderate the effect of weak explicit knowledge sharing on developer's contribution level.
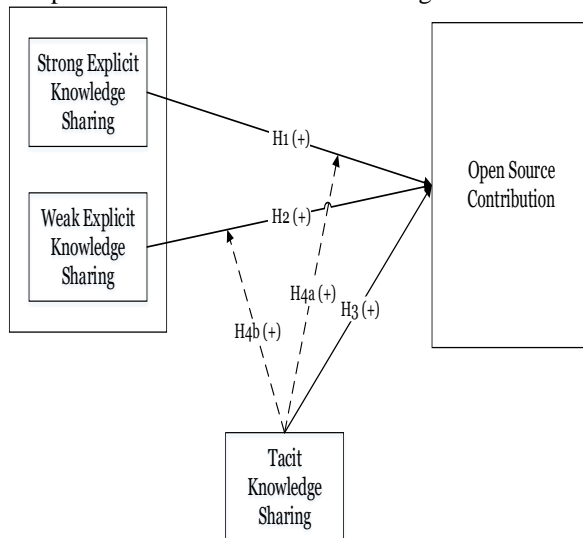We present our theoretical model in Figure 1.

Figure 1 Theoretical model for the effect of knowledge sharing on knowledge contribution

4. Research Context and Data

As discussed earlier the context of our research pertains to open source communities – Github and Gitter. We wanted to understand how these two platforms operate and how their respective interaction plays the role of a catalyst in facilitating knowledge sharing, which in turn impacts the developer's contribution levels.

4.1. GitHub platform

The GitHub platform is the largest OSS community in the world with more than 4 million software developers involved [6]. The platform hosts the programming code for more than 9 million OSS projects. In the last decade, it has become synonymous with online collaboration and real-time code sharing between developers. The platform facilitates this by providing a plethora of features that are directed towards distributed version control of projects.

On the platform, each project is stored as either a repository or a set of repositories. The Wikipedia page of the repository provides an overview of the goal and direction of the project. Each change made by a developer is recorded as a *commit*, which helps keep track of which team member made what changes. This tracking mechanism also serves as the version control as it provides the project maintenance team with the option of rolling back to previous versions.

The overarching goal of the platform is to facilitate individuals to collaborate on OSS projects whilst enabling the project maintenance team to have proper version control. As such multiple other features like 'watch', 'release' and 'push' amongst others have been developed. We discuss these features in greater detail in the section below. The detailed information of user activity on the platform can be obtained using GitHub's API.

The GitHub data we collected consists of monthly user level data for 3,695 developers over the span of 2017. As such, we have 44,340 observations in our dataset. We present the summary statistics in Table 1. The attributes pertain to the various activities that a developer performs on the platform. The variables in our dataset follow exponential distribution. The exponential distribution of variables can be attributed to nature of data over Internet (Open source is an online community) where few individuals perform majority of activities and majority of individuals are only watching and do not participate in any given time period. Therefore, in such data, a large number of zeros (min. value) is expected. We observe that there is a very high level of heterogeneity within the user's activity levels, for example, some users have very high number of pull requests - log (pull requests) (max. value = 5.897), whereas others hardly have any (min. value = 0). We also computed correlation for the explanatory variables and observe that none of the bivariate correlations are greater than 0.4. These low correlations among the variables suggest that they are not related to each other and contribute uniquely to the model in explaining the dependent variable.

4.2. Gitter platform

"Gitter chat" or commonly known as Gitter is a chat room based mobile and web application. It provides real-time computer mediated communication to the developers. The chat rooms are created based on technology topics such as Python programming, R Programming, etc., and based on open source repositories such as Scikit-learn. When the chat room is created for an open source project, the creator of the chat room can also link the events of the open source project with the chat application. Such events include activities related to opening or closing of an issue, commits, etc.
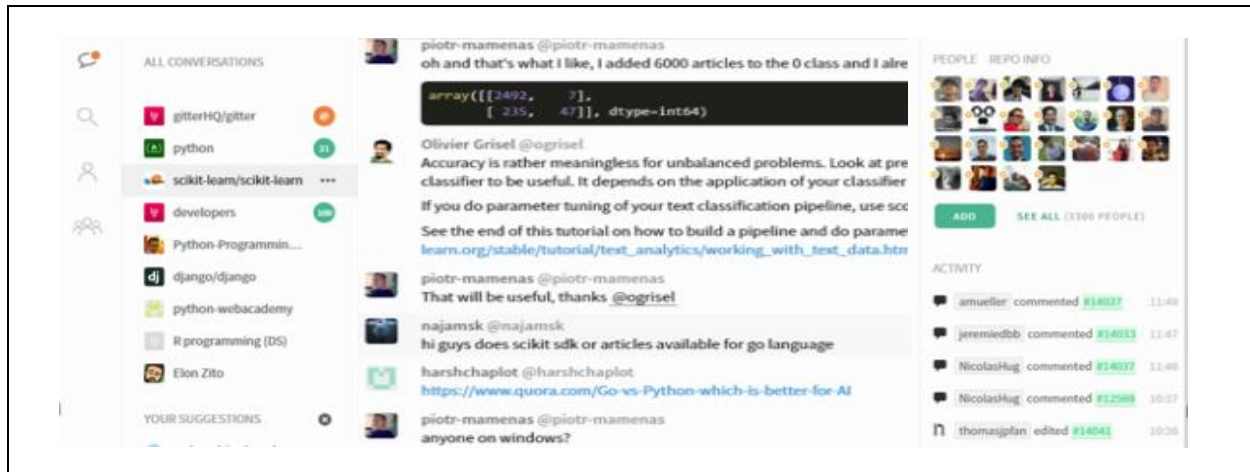
Figure 2 Screen shot of the Gitter platform.

The screenshot in Figure 2 shows the Gitter web application. The left panel in the screenshot shows the chat room. An individual can select a chat room and join it. After joining the chat room, the individual can see recent chat messages in the central window. In this window, an individual having a query can ask questions and reply to other participants. The right top window shows the available participant active in the chat. The bottom right window shows the recent activities from attached GitHub project repository. All these features are also available for the mobile. The Gitter data is collected from Kaggle website. We analyzed data for year 2017. We took a random sample of 3,695 users and mapped with their activities on GitHub platform during the same time period. These activities have been explained in the following section.

5. Measures

As presented earlier, our study makes use of developer's activities on two of the biggest open source communities, namely GitHub and Gitter. We have been able to collect data that not only keeps track of a developer's activities on the GitHub platform but also follows their chat activity on the Gitter platform. This provides us with the opportunity to not only have a measure for their open source contribution, but also evaluate the different avenues that they may have used to share and obtain knowledge.

We identify the total number of 'commits' made by a developer to be the representation of their overall contribution on GitHub. Commits has been used in multiple studies in the past to be a robust measure of contribution on the platform [6].

To measure different mediums of knowledge sharing we make use of the following variables: 'Wiki', 'isChat' and 'forum'. The total number of Wikipedia edits made by a developer is captured by the variable Wiki and is used as measure of their strong explicit knowledge sharing in our study. Whereas, the total number of questions raised along with the total number of responses provided by the developer is captured by 'forum' variable and used as the measure for the individual's weak explicit knowledge sharing. In addition, the individual's activity on Gitter platform is measured as whether or not they have sent any chat messages in a particular month, this is captured by the isChat variable in our study and represents the tacit knowledge sharing in our study.

Other variables that we use in our study as control variables are listed below:
• Create: This is used to track how many new repositories or branches that have been created by the developer in each month.
• Delete: This is used to track how many repositories or branches that have been deleted by the developer in each month.
• Fork: This variable is used to track how many times in a month a developer forks a repository to create a personal copy for themselves. They can make their own changes to the said copy without affecting the original project.
• Member: This is used to track how many times a developer has been added to projects as a collaborator or removed from projects.
• Pull Request: A pull request consists of one or more commits. A pull request is triggered when a pull request is assigned, unassigned, labeled, unlabeled, opened, edited, closed, reopened, synchronized, a pull request review is requested, or a review request is removed. This variable tracks the number of pull requests submitted by a developer in a month.
• Pull Request Review Comment: After raising pull request, the team members in an OSS project can review the code changes and comment on that if required. These comment events are recorded as 'Pull

Request Review Comment.' Here we track the number of such comments made by the developer in a month.

• Push: This variable represents number of push events which are triggered when code is pushed to a repository branch. Branch pushes and repository tag pushes also trigger web-hook push events. Hence, the variable tracks the total number of pushes made by a developer in each month.

• Release: This variable tracks the number of releases made by a developer in a month. A release is triggered by key developers from the repository when a new release of the repository is made available for download.

• Watch: This variable tracks the number of times in a month a developer marks repository to watch for future announcements.

• Public: This variable tracks the number of times in a month, a developer changes a repository from private to public

The variables in our dataset follow exponential distribution. The exponential distribution of variables can be attributed to nature of data over the Internet (Open source is an online community) where few individuals perform majority of activities and majority of individuals are only watching and do not participate in any given time period. Therefore, in such data, a large number of zeros are expected.

To transform the exponential distributed variable to a linear distribution, we take logarithm transformation of our variables. The log transformation is performed following [42] wherein we take logarithm of variable+ constant, with constant=1 instead of directly taking logarithm of variable. The approach by Fletcher et al. [42] addresses the issue of taking logarithm of '0', which is mathematically not defined. We have used constant =1 with logarithm base as 'e'. We present the summary statistics of the variables in Table 1.

Variable	Observations	Mean	Std. Dev.	Min	Max
log (Commit)	44,340	0.004	0.074	0	2.833
log (Wiki)	44,340	0.004	0.082	0	4.533
isChat	44,340	0.098	0.298	0	1.000
log (Forum)	44,340	0.067	0.372	0	5.710
log (Create)	44,340	0.244	0.634	0	5.050
log (Delete)	44,340	0.029	0.213	0	5.247
log (Fork)	44,340	0.075	0.315	0	5.878
log (Member)	44,340	0.010	0.106	0	2.303
log (Public)	44,340	0.003	0.058	0	2.890
log (Pull Request)	44,340	0.073	0.385	0	5.897
log (Pull Request Review Comment)	44,340	0.006	0.113	0	4.407
log (Push)	44,340	0.326	0.882	0	8.113
log (Release)	44,340	0.003	0.066	0	4.625
log (Watch)	44,340	0.146	0.473	0	5.624

Table 1 Summary statistics

DV: log (Commit)	(1) FE	(2) FE with Controls	(3) RE	(4) RE with Controls
log (Wiki) (H1)	0.047***	0.040***	0.055***	0.043***
	(0.005)	(0.005)	(0.004)	(0.004)
isChat (H3)	0.001	-0.003**	0.001	-0.003***
	(0.001)	(0.001)	(0.001)	(0.001)
log(Forum) (H2)	0.037***	0.024***	0.041***	0.024***
	(0.001)	(0.002)	(0.001)	(0.001)
isChat # log (Wiki) (H4a)	0.131***	0.127***	0.129***	0.129***
	(0.020)	(0.020)	(0.020)	(0.019)
isChat # log (Forum) (H4b)	0.005*	0.005**	0.005**	0.006***
	(0.002)	(0.002)	(0.002)	(0.002)
log (Create)		0.004***		0.003***
		(0.001)		(0.001)
log (Delete)		0.007***		0.0154***
		(0.002)		(0.002)
log (Fork)		0.008***		0.006***
		(0.001)		(0.001)
log (Member)		0.017***		0.016***
		(0.004)		(0.003)
log (Public)		-0.024***		-0.015**
		(0.007)		(0.006)
log (Pull Request)		0.004**		0.004***
		(0.002)		(0.001)
log (Pull Request Review Comment)		0.042***		0.0347***
		(0.004)		(0.004)
log (Push)		0.002***		0.001
		(0.001)		(0.001)
log (Release)		0.040***		0.049***
		(0.006)		(0.005)
log (Watch)		0.005***		0.006***
		(0.0012)		(0.001)
Constant	0.002***	-0.001***	0.001**	-0.001*
	(0.0004)	(0.0004)	(0.001)	(0.0005)
Observations	44,340	44,340	44,340	44,340
R-squared	0.028	0.041		
Number of User	3,695	3,695	3,695	3,695

Standard errors in parentheses

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

Table 2 Regression result – Panel Model

6. Results

To test our hypothesis, we conduct a panel linear regression and present our results in Table 2. We use fixed-effects models to control for omitted time-invariant developer characteristics. We have also reported results from random effect model. We can see from Table 2, fixed-effect model (column 1), that the strong explicit knowledge sharing has positive effect on contribution, which supports our hypothesis H1 ($p<0.01$). Similarly, we also find support for our hypothesis H2 ($p<0.01$) which states that the weak

explicit knowledge sharing has positive effect on the contribution. However, we do not find any support for our hypothesis H3 which states that the tacit knowledge transfer positively contributes to contribution. We have included a plausible explanation for this result in our discussion section.

We also test for the moderating effect of tacit knowledge sharing on the relationship between explicit knowledge sharing and contribution. We find that hypotheses – H4a (p<0.01) and H4b (p<0.1) are supported. Tacit knowledge sharing positively moderates the relationships of both strong explicit knowledge sharing and weak explicit knowledge sharing on contribution.

7. Discussion

We find clear support of all the hypotheses except hypothesis H3. Hypothesis H3 states that the tacit knowledge sharing has positive effect on the contribution. We can see that the tacit knowledge – log(isChat) in fixed effects model in Table 2 is not significant while it is significant in fixed effect model with controls.   One plausible explanation for the rejection of this hypothesis is that the tacit knowledge sharing helps to clarify the explicit knowledge sharing and does not contribute on its own. This might explain the discrepancy between the significance of tacit knowledge sharing in the two fixed effects models. Moreover, in our current analysis, we have not included project level data. The chat rooms on the Gitter platform either are related to an open source project repository or are created by individuals based on common interests. Some of the chat rooms do not have a relationship with an active open source project. In our future analysis, we plan to select chat rooms that are dedicated to a GitHub repository to control for the above confounding effect.

8. Conclusion

In this paper, we elaborate on different knowledge sharing mediums community and examine the effect of explicit and tacit knowledge sharing on the contribution in open source community. We extend the taxonomy of knowledge sharing by classifying the explicit knowledge sharing in two categories – strong explicit knowledge sharing and weak explicit knowledge sharing. The results demonstrate that both types of external knowledge sharing positively influence contribution in open source project development. Moreover, the tacit knowledge sharing between individuals positively moderates the above relationship. Our research has few limitations. First,

we use commit as a signal for contribution. While this may be a good signal for most scenarios, there might be instances where the developer might commit without including any information. Second, we do not control for the amount or type of information exchanged in explicit and tacit knowledge sharing mediums. We plan to conduct content analysis to quantify and categorize information included in each commit for future research.

Overall, our paper contributes to the theoretical understanding of knowledge sharing by demonstrating that the tacit knowledge sharing improves the effectiveness of explicit knowledge sharing. Moreover, our paper also contributes to practice of open source development by demonstrating the need for tools and techniques required for more effective tacit knowledge sharing.

9. References

[1] Fitzgerald, Brian. 2006. "The Transformation of Open Source Software," *MIS Quarterly*, (30: 3).
[2] AlMarzouq, M., Zheng, L., Rong, G., & Grover, V. 2005. "Open Source: Concepts, Benefits, and Challenges," *Communications of the Association for Information Systems*, 16
[3] Mehra, Amit and Mookerjee, Vijay M. 2012. "Human Capital Development for Programmers Using Open Source Software," *MIS Quarterly*, (36: 1)
[4] Watson, Richard T.; Wynn, Donald; and Boudreau, Marie-Claude 2008. "JBOSS: The Evolution of Professional Open Source Software," *MIS Quarterly Executive*, Vol. 4(3).
[5] Kavanagh, J., 2004. "Resistance as Motivation for Innovation: Open Source Software." *Communications of the Association for Information Systems*, 13.
[6] Moqri, M., Mei, X., Qiu, L. and Bandyopadhyay, S., 2018. "Effect of "Following" on Contributions to Open Source Communities." *Journal of Management Information Systems*, 35(4).
[7] Barney, J. 1991. "Firm Resources and Sustained Competitive Advantage," *Journal of management*, 17(1).
[8] Bitzer, J., and Geishecker, I. 2010. "Who Contributes Voluntarily to Oss? An Investigation among German It Employees," *Research policy* 39(1).
[9] Bitzer, J., Schrettl, W., and Schröder, P. J. 2007. "Intrinsic Motivation in Open Source Software Development," *Journal of comparative economics* (35:1), pp. 160-169.
[10] Héritier, A., and Lehmkuhl, D. 2008. "The Shadow of Hierarchy and New Modes of Governance," *Journal of public policy* (28:1), pp. 1-17.
[11] Lee, J.-N. 2001. "The Impact of Knowledge Sharing, Organizational Capability and Partnership Quality on IS Outsourcing Success," *Information & management* (38:5), pp. 323-335.
[12] Liu, Y., and Phillips, J. S. 2011. "Examining the Antecedents of Knowledge Sharing in Facilitating Team Innovativeness from a Multilevel Perspective,"

International Journal of Information Management (31:1), pp. 44-52.

[13] Ma, M., and Agarwal, R. 2007. "Through a Glass Darkly: Information Technology Design, Identity Verification, and Knowledge Contribution in Online Communities," *Information systems research* (18:1), pp. 42-67.

[14] Bonaccorsi, A.; and Rossi, C., 2006. "Comparing motivations of individual programmers and firms to take part in the open source movement: From community to business," *Knowledge,Technology & Policy*, 18, 4 (2006), 40–64.

[15] Chiu, C.; Hsu, M.; and Wang, E.T. 2006. "Understanding knowledge sharing in virtual communities: an integration of social capital and social cognitive theories," *Decision Support Systems*, 42 (3), 1872–1888.

[16] Goes, P.B., Lin, M.; and Yeung, C.-M.A., 2014. "Popularity effect in user-generated content: evidence from online product reviews," *Information Systems Research*, 25, 2 (2014), 222–238.

[17] Kankanhalli, A.; Tan, Bernard C.Y.; and Wei, K.-K., 2005. "Contributing knowledge to electronic knowledge repositories: An empirical investigation," *MIS Quarterly*, 113–143.

[18] Khansa, L.; Ma, X.; Liginlal, D.; and Kim, S.S., 2015. "Understanding members' active participation in online question-and-answer communities: A theory and empirical analysis," *Journal of Management Information Systems*, 32, 2 (2015), 162–203.

[19] Okoli, C.; and Oh, W., 2007. "Investigating recognition-based performance in an open content community: A social capital perspective," *Information & Management*, 44, 3 (2007), 240–252.

[20] Wasko, M.; and Faraj, S., 2000. "It is what one does: Why people participate and help others in electronic communities of practice," *The Journal of Strategic Information Systems* 9, 2 (2000), 155–173.

[21] Wasko, M.; and Faraj, S., 2005. "Why should I share? Examining social capital and knowledge contribution in electronic networks of practice," *MIS Quarterly*, 29, 1, (2005), 35–57.

[22] Zhang, X.; and Zhu, F., 2011. "Group size and incentives to contribute: A natural experiment at Chinese Wikipedia," *The American Economic Review*, 101, 4 (2011), 1601-1615.

[23] Zhao, L.; Detlor, B.; and Connelly, C.E., 2016. "Sharing knowledge in social Q&A sites: The unintended consequences of extrinsic motivation," *Journal of Management Information Systems*, 33, 1 (2016), 70–100.

[24] P. Vixie, in: C. Dibona, S. Ockman, M. Stone (Eds.), 1999. "Software Engineering. In Open Sources: Voices from the Open Source Revolution," *O'Reilly*, Sebastopol, 1999.

[25] G. Hertel, S. Niedner, S. Herrmann, 2003. "Motivation of software developers in open source projects: an Internet-based survey of contributors to the Linux kernel," *Research Policy*, 32, 2003.

[26] A. Bonaccorsi, C. Rossi, 2003. "Why open source software can succeed," *Research Policy* 32 (7), 2003, pp. 1243–1258.

[27] Dossi, G. 1988. "Sources, Procedures, and Microeconomic Effects of Innovation," *Journal of Economic Literature*, 26 pp. 1120–1171.

[28] Fleming, L., D.M. Waguespack. 2007. "Brokerage, Boundary Spanning, and Leadership in Open Innovation Communities," *Organization Science*, 18(2) pp 165–180.

[29] Nuvolari, A., 2005. "Open Source Software Development: Some Historical Perspectives," *First Monday*, 10(10).

[30] Singh, P.V. 2007. "Open Source Software and the Small World Phenomenon: An empirical Investigation of Macro Level Network Properties on Project Success," *Proceedings of the 28th ICIS*.

[31] Michael, P., 1958. "Personal knowledge. Towards a post-critical philosophy," With a New Foreword by Mary Jo Nye.-2015.-284 p.

[32] Goffin, K. and Koners, U., 2011. "Tacit knowledge, lessons learnt, and new product development," *Journal of Product Innovation Management*, 28(2), pp.300-318.

[33] S. Tamer Cavusgil, Roger J. Calantone, Yushan Zhao, 2003. "Tacit knowledge transfer and firm innovation capability," *Journal of Business & Industrial Marketing*, Vol. 18 Issue: 1, pp.6-21, https://doi.org/10.1108/08858620310458615

[34] Nonaka, Ikujiro. "A dynamic theory of organizational knowledge creation." *Organization science* 5.1 (1994): 14-37.

[35] Polanyi, Michael. "The logic of tacit inference." *Philosophy*41.155 (1966): 1-18.

[36] Collins, Harry. *Tacit and explicit knowledge*. University of Chicago Press, 2010.

[37] Reagans, Ray, and Bill McEvily. "Network structure and knowledge transfer: The effects of cohesion and range," *Administrative science quarterly* 48.2 (2003): 240-267.

[38] Panahi, Sirous, Jason Watson, and Helen Partridge. "Social media and tacit knowledge sharing: Developing a conceptual model," *World academy of science, engineering and technology* 64 (2012): 1095-1102.

[39] Hamel, Gary. "Competition for competence and interpartner learning within international strategic alliances." *Strategic management journal* 12.S1 (1991): 83-103.

[40] Chen, Yongmin., 2005. "Vertical disintegration," *Journal of Economics & Management Strategy* 14.1 (2005): 209-229.

[41] Nobeoka, Kentaro. "Inter-project learning in new product development." *Academy of Management Proceedings*. Vol. 1995. No. 1. Briarcliff Manor, NY 10510: Academy of Management, 1995.

[42] Fletcher, D., MacKenzie, D. and Villouta, E., 2005. "Modelling skewed data with many zeros: a simple approach combining ordinary and logistic regression," *Environmental and ecological statistics*, 12(1), pp.45-54.

[43] Marwick, Alan D. "Knowledge management technology." IBM systems journal 40, no. 4 (2001): 814-830.

[44] Peddibhotla, Naren B., and Mani R. Subramani. "Contributing to public document repositories: A critical mass theory perspective." *Organization Studies* 28, no. 3 (2007): 327-346.