12-12-2020

Disruptions in Open Source Platforms: A Natural Experiment with COVID-19 Pandemic

Dr. Babu Veeresh Thummadi

Disruptions in Open Source Platforms: A Natural Experiment with COVID-19 Pandemic

(Completed Research Paper)

Dr. Babu Veeresh Thummadi
Lero, The Science Foundation Research Centre for Software,
Cork University Business School, University College Cork, Ireland.
vthummadi@ucc.ie

ABSTRACT

Disruptions to open source software platforms can occur for various reasons such as natural disasters, cyber-attacks etc., that are beyond the control of open source workers. As open source work relies primarily on voluntary contributions, any type of disruptions to the work routines of open source workers can lead to loss of productivity and cause delays in software production and releases. Currently, we do not have sufficient understanding about worker resilience or the effects of disruption on open source development. To understand the effects of disruption, data has been collected on open source production activities during the COVID-19 pandemic, an exemplar event, from January 1st 2020 to 7th April, 2020 from two different locations that had different lockdown policies. The results highlight the importance of disruptions in open source production environments and project management practises during crisis situations.

Keywords

Open source, COVID-19, disruptions, online communities.

INTRODUCTION

Open source software is a software that is developed by individuals or organizations with contributions from voluntary workers through web-based platforms (Von Hippel and Von Krogh 2003) such as Github, Bitbucket etc., and have challenged the traditional ways of developing software and project management activities within organizations. Unlike, proprietary software, open source can be rich in kind due to its wide-scale availability and enhancements provided by altruistic people participating in software development activities (Puranam et al. 2014). Hence organizations are adopting both open source software and open source development practices into organizations through inner-sourcing practices (Stol and Fitzgerald 2014).

Open source platforms are useful for entrepreneurs and organizations for tapping the workforce and IT skills that are beyond the reach of the organizational boundaries, however, open source platforms are also susceptible to disruptions which can lead to potential problems in coordination, production releases and open source project management (Lindberg et al. 2016). We define here disruptions in open source platforms as triggering events that are "sufficiently jarring that it cannot be ignored" (Morgeson and DeRue 2006, p 272). One key feature that distinguishes disruptions from other forms of "interruptions" or "distractions" (such as a simple email or pull request or a comment) is that they are sufficiently large and unexpected and have the potential for creating chaos and uncertainty in open source platforms (Stoverink et al. 2020).

This is problematic as disruptions can create challenges to users that routinely build software through these open-source platforms. Disruptions can bear heavily on open source processes, software release cycles and voluntary contributions (Stoverink et al. 2020) and despite over decades of research on open source platforms there are not many studies that have studied the phenomena of disruption and how it affects open source activities i.e., what sort of activities get highly disrupted and which activities are least affected and for how long? Hence, we can ask:

> *RQ1: What is the impact of disruption on open-source activities? And how long does it take for workers to reverse this effect?*

Recent studies on digital platforms (Choudhury et al. 2020; Zhang and Zhu 2011), banking (Dong et al. 2018) have used “difference-in-difference” estimation technique to test causality by using natural experiments. To this end, our research design follows this logic to understand how disruption affects open source platform activities. To this end, data was collected from a popular open source site Github (www.github.com) which was established in the year 2008 and has over 40 million users from all over the world in 2020 (Gousios 2013). We chose a major event dubbed “COVID-19 pandemic” (a pandemic that was caused due to a novel coronavirus outbreak in Wuhan in December 2019) as an exemplar of disruption in open source platforms as open source workers are subject to different working conditions (Liu et al. 2020). As COVID-19 pandemic affected different regions to different degrees, governments have imposed different types of lockdowns and thus allowing us to understand how different lockdown policies (thereby different open source working conditions) can affect the opensource production activities of individual workers.

For setting up a quasi-natural experiment, data was chosen on activities of users from two locations that have faced different type of disruptions during the COVID-19 pandemic, i.e., one with no lockdown and one with full lockdown during the same period (Liu et al. 2020). The analysis shows that contributions decrease by 19% due to disruption, however, the effect was much smaller in entrepreneurial activities. At the same time, the effect of disruption was observed to be persistent for a long time.

The paper is organized as follows. First literature review is provided on open source platforms and disruptions. Then research design is discussed. Next key findings are discussed. The paper ends with discussion, conclusion and suggest possible avenues for future research.

BACKGROUND

Research on open source platforms spans across multiple disciplines in the fields of information systems, management, and organization science and has been touted under different labels such as open-source science, private-collective innovation model, community-based innovation, open innovation, Open Source Software Development (OSSD), virtual organizations etc., to name a few (Foss et al. 2016).

Open source platforms rely on the concept of developing software through volunteers via web-based hosting services (ibid). In line with this view, projects developed in open source platforms depend on the volunteers to carry out tasks such as bug reporting, code development, project management and develop code collectively through patches and software releases. Researchers have made several important contributions to unpack the inner workings of how open source projects work and how they are governed and what enables the projects to remain solvent (Von Hippel and Von Krogh 2003). Studies pinpoint the reasons for success (Crowston et al. 2003; Crowston and Howison 2005; Crowston et al. 2006), and how trust between developers enhances open source effectiveness (Stewart and Gosain 2006), and the role of diversity in community engagement (Daniel et al. 2013), and communication strategies for growth of open source projects (Foss et al. 2016).

Surprisingly, despite over decades of research on open source platforms there have not been many studies that have studied the phenomena of disruption and how it affects open source activities i.e., what sort of activities get highly disrupted and which activities are least affected and for how long? Some studies have been conducted in the area of software development for understanding the role of “interruptions” (a superset of disruptions) to shed light on the effective strategies that organizations can utilize in reducing the negative effects of interruptions (Sykes 2011). At the same time, a recent study on online communities showcase the importance of self-organizing during crisis situations (Nan and Lu 2014) and suggest that more work is needed in unpacking the nuances of open source production process such as platform disruptions due to exogeneous shocks (see Fisher 2019 and Faraj et al 2011, 2016 for more details regarding the differences between open source and online communities)(Faraj et al. 2011; Faraj et al. 2016; Fisher 2019). Hence Nan and Liu (2014) note that: “online community research, being relatively new, is still searching for fruitful angles to grasp the fundamental difference between online communities and traditional forms of communication and collaboration (Majchrzak 2009). Since self-organization in online communities manifest organizing dynamics distinct from those in traditional communication and collaboration platforms, it provides a promising new angle to rethink the key constructs of existing online community research: motivation, structural mechanism, and governance”(Nan and Lu 2014, p 1153).

Understanding the effects of disruptions on open-source platforms is important for several reasons. In today's environment, disruptions to open source platforms are increasingly becoming common due to cyber-attacks, power outages, natural disasters, censorships etc., (Zhang and Zhu 2011). All these type of disruptions can bear heavily on open source software processes, production release and voluntary contributions (Stoverink et al. 2020). There have not been many studies that have studied the phenomena of disruption and how it affects open source activities i.e., what sort of activities get highly disrupted and which activities are least affected and for how long? As we know that most of the firms are dependent on open source platform disruptions can be problematic for organizations (Nagle 2019).

Second, disruption can impair team functioning and cause delays in the open source production process, especially, when open source team or an individual member is faced with a tricky situation or a puzzle for which there are no clear solutions (Stoverink et al. 2020). As most of the open source development is self-organized (Foss et al. 2016), the product owner or manager has to monitor and prepare strategies for coping up with the disruptive events to reduce delays in production. To this end, product owners managing open source need to understand the type of disruptive events that can be expected in open source and how to effectively manage their teams by understanding the potential impact of disruption on workers (Morgeson and DeRue 2006). For example, a recent news report finds that, when a software programmer[1] pulled out the code from a popular package called Node Package Manager (NPM) due to legal issues it resulted in disruption across thousands of projects that use this software code. The CEO of the company had to take unprecedented steps to bring the situation under control. This exemplifies the heightened challenges in managing open source development during disruptions.

Third, when platforms are disrupted, open source teams need to develop resilience to overcome the situation (Stoverink et al. 2020). This can be challenging for many open source software teams due to lack of training and project management experience. In conclusion, disruption(s) is an understudied topic in open source literature and through this study we would like to shed light in this area. Next, we discuss our research design.

RESEARCH DESIGN

In this study we followed a research design that is traditionally followed in observational studies in information systems (Horton and Tambe 2015). We illustrate our research design with a figure to give an overview of the process followed (see Figure 1). First the research questions were formulated and an quasi-experiment was set up and the data was collected in the next step and then exploratory data analysis was carried out. Finally, the results were interpreted.

[1] See https://www.theregister.co.uk/2016/03/23/npm_left_pad_chaos/

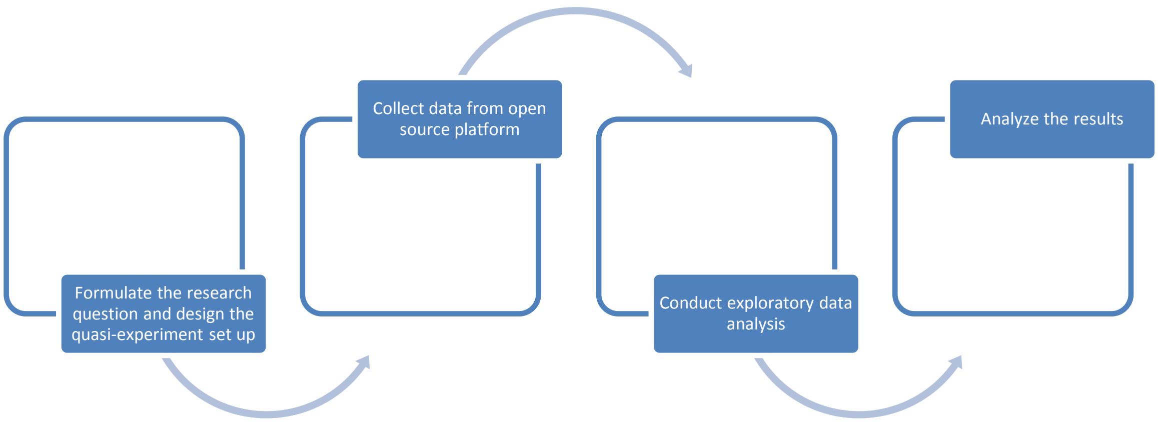

Figure 1: Research design

Data collection

Data was collected from two different sources. First, data was collected from a popular open source platform, Github, which has over 100 million projects and 40 million users all over the world as of May 2020 (see https://en.wikipedia.org/wiki/GitHub). Specifically, data was collected from GHTorrent database which is widely used by the researchers for studying open source platform (for example see Medappa and Srivastava 2019 for more details) (Medappa and Srivastava 2019).

In order to understand the effects of disruption, the COVID-19 pandemic has been chosen as a setting as this allows to conduct a natural experiment across two different locations that were disproportionately disrupted by the government policies (i.e., in one location there was full lockdown and causing disruption and in another location had no lockdown and hence less disruption) (Choudhury et al. 2020; Wu et al. 2020). To this end, two locations were chosen, namely Wuhan, China and Hong Kong, China for estimating the disruption effects on open source workers. Wuhan was chosen as this location was the worst hit of the COVID-19 pandemic and where the lockdown was enforced early on from January 23rd, 2020 until April 8th, 2020[2]. Hong Kong was chosen as a control group as this location (some people think Hong Kong is a country but it is part of China and referred to as special administrative region) (Kam Ng 2008) adopted a different strategy and did not enforce lockdown but rather used surveillance technologies, contact tracing and quarantine methods to fight the COVID-19 pandemic. Both Wuhan and Hong Kong almost have similar populations, Wuhan has roughly around 11 million and Hong Kong has roughly about 7 million people[3], and both are metropolitan cities, and thus allowing for a quasi-natural experiment set up to understand the effects of disruption on open source worker productivity.

GHTorrent dataset provides details of the individual users, location and other demographic details. These details are being leveraged to identify the users from Wuhan and HK. Using fuzzy string-matching techniques, 3341 users (2637

---

[2] See https://en.wikipedia.org/wiki/COVID-19_pandemic_lockdown_in_Hubei

[3] See https://en.wikipedia.org/wiki/List_of_cities_in_China_by_population

from Hong Kong and 704 from Wuhan) were identified that would be ideal for the study. Users were eliminated who have listed multiple locations such as "Hong Kong/New York". The user activity data was collected between January 1st, 2020 and April 7th, 2020 on open source activities of workers (Gousios 2013). One limitation of the GHTorrent dataset is that many users do not display their locations and the database of the users sometimes are not updated, thus leading to underrepresentation of all the users in the location, and could lead to selection bias in the sample. As this study, compares the individual productivity across locations, any misrepresentations should likely even out across locations, however this should be considered as a limitation to the study.

A second dataset on COVID-19 infections was collected in the locations to account for the possible exposure of the disease to the user[4]. This dataset is provided by Johns Hopkins University Center for Systems Science and Engineering (JHU CCSE). Wuhan and Hong Kong drastically differed in terms of the growth of the disease, hence this is used as a control variable for possible effects of the COVID-19 pandemic. Now the variables in the study are discussed in detail.

Dependent variables: For the study, a count of all the Github activities (for example forking, creating, deleting etc.,) of worker i at time t for measuring the All Contributions$_{it}$. This type of measures have been used in earlier studies (who has used similar type of proxy measures, see Belenzon and Schankerman 2015 and Zhang and Zhu 2011 for more details) (Belenzon and Schankerman 2015; Zhang and Zhu 2011). *All contributions* $_{it}$ is measured for individuals on a weekly basis by accounting for the exact times in China (data on GHTorrent shows the timestamps in UTC format and hence were transformed to time in China). This measure is log transformed and is coded as log (1+x), where x refers to the total number of open source contributions. In addition, other dependent variables have been generated which are grouped by the activity types, for example, *Entrepreneurship*$_{it}$ (is grouped as sum of all the create events), *Personal learning*$_{it}$ (sum of all fork events and watch events) and the rest all activities are grouped under *Project management*$_{it}$ variable (for example push event, pull request event etc.,).

Treatment variables: *Location*$_i$ variable is coded as 1 if the individual i location is in Wuhan and 0 for Hong Kong. The treatment variable *After*$_t$ is coded as 1 for 5 to 14 weeks of year 2020 (the lockdown started in Wuhan on January 23rd, 2020 i.e., 4th week of the year, as we cannot determine if this week belonged to post or pre-treatment we dropped these observations) and first three weeks of January 2020 are coded as 0 (i.e., for pre-treatment). The interaction term is coded based on the prior two variables. It is a multiplicative product of *After*$_t$ * *Location* $_i$.

Controls: Tenure $_{it}$, *(Tenure)^2* $_{it}$ and *COVID-19 infections* $_t$ variables are used as controls. *Tenure* $_{it}$ is calculated as the total number of weeks of the worker in open source worker. *(Tenure)^2* $_{it}$ term is included for possible curvilinear effects. *Tenure* $_{it}$ is divided by 100 and *(Tenure)^2* $_{it}$ by 10000 for presentation of the findings. *COVID-19 infections* $_{it}$ is the total number of COVID-19 infections in prior week in the local region where the user is present. Unfortunately, the Humanitarian data used in the dataset is provided for Hubei province instead of Wuhan, nonetheless, most of the cases in the January and February originated from Wuhan hence Hubei should be good proxy for the COVID-19 exposure. At the same time, data is provided from 23rd January, 2020. Hence, we manually calculated the possible number of cases in Wuhan by using an epidemic calculator for the first three weeks. We chose R_0=10 and R_t= 9.3 and estimated the total cases by taking December 1st as the origin of the epidemic in Wuhan (Liu et al. 2020). For, Hong Kong, there were no cases in prior weeks of January 23rd and thus we coded them as 0. This variable is divided by 100000 for presentation of the findings.

Data analysis

To estimate the effects of the treatment, difference in difference estimator technique is used (Bertrand et al. 2004; Wooldridge 2007). As the shock due to pandemic was natural and unexpected, the users in the platform can be thought of as if randomly chosen for disruption treatment. Both OLS and fixed effects regressions are carried out as the dependent variable is a count variable. The dependent variable is transformed as log (1+x) and for the both the OLS and fixed effects regression.

---

[4]See https://data.humdata.org/dataset/novel-coronavirus-2019-ncov-cases

Summary statistics

In table 1, descriptive statistics are reported. As mentioned before the total number of observations in the dataset is 43433 and there are 3341 individuals in total from both treated and untreated locations. We can see the standard deviation for COVID-19 infections is high which is expected considering the massive outbreak in Wuhan. In table 2 the correlation between variables are provided. We can see that tenure is positively correlated with the dependent variables which was expected.

Table 1: Descriptive statistics

Variables	Mean	SD	Min	Max
All contributions $_{it}$	0.23	0.69	0.00	7.35
Entrepreneurship $_{it}$	0.04	0.25	0.00	5.71
Personal learning $_{it}$	0.10	0.36	0.00	4.91
Project management $_{it}$	0.15	0.59	0.00	7.18
Location $_{i}$	0.21	0.41	0.00	1.00
After $_{t}$	0.77	0.42	0.00	1.00
Tenure $_{it}$	3.44	0.90	1.92	6.20
(Tenure) $_{it}$ ^2	12.66	6.66	3.69	38.44
COVID-19 infections $_{it}$	0.90	2.11	0.00	6.78

N=43433

Table 2: Correlation for all variables

	1	2	3	4	5	6	7	8	9
1. All contributions $_{it}$	1								
2.Entrepreneurship $_{it}$	0.66^{***}	1							
3. Personal learning $_{it}$	0.64^{***}	0.22^{***}	1						
4. Project management $_{it}$	0.89^{***}	0.68^{***}	0.28^{***}	1					
5. Location $_{i}$	-0.00	-0.01^{**}	0.06^{***}	-0.03^{***}	1				
6. After $_{t}$	0.01^{**}	0.01^{**}	0.01	0.01^{**}	0.00	1			
7.Tenure $_{it}$	0.13^{***}	0.06^{***}	0.10^{***}	0.10^{***}	-0.07^{***}	0.04^{***}	1		
8. (Tenure) $_{it}$ ^2	0.13^{***}	0.06^{***}	0.10^{***}	0.10^{***}	-0.08^{***}	0.03^{***}	0.99^{***}	1	
9. COVID-19 infections $_{it}$	0.01	-0.00	0.05^{***}	-0.02^{***}	0.76^{***}	0.23^{***}	-0.04^{***}	-0.05^{***}	1

Note: + p<0.10, * p<0.05, ** p<0.01, *** p<0.001

RESULTS

In this section results are reported. First, a model free evidence of the effect of treatment in the treated group vs. untreated group is shown visually (see Figure 2). Both locations show parallel trends before the treatment and we can see in the 4th week of January 2020 there is drastic reduction in the overall contributions in both locations, likely due to the lockdown announcement on January 23rd, 2020, though the fall is not as sleep in the untreated location (parallel trends have been observed for other dependent variables as well but have not been displayed due to space limitations). As suggested by Mora and Reggio (2012) additional test has been carried out for confirming the parallel trends (Mora and Reggio 2012).

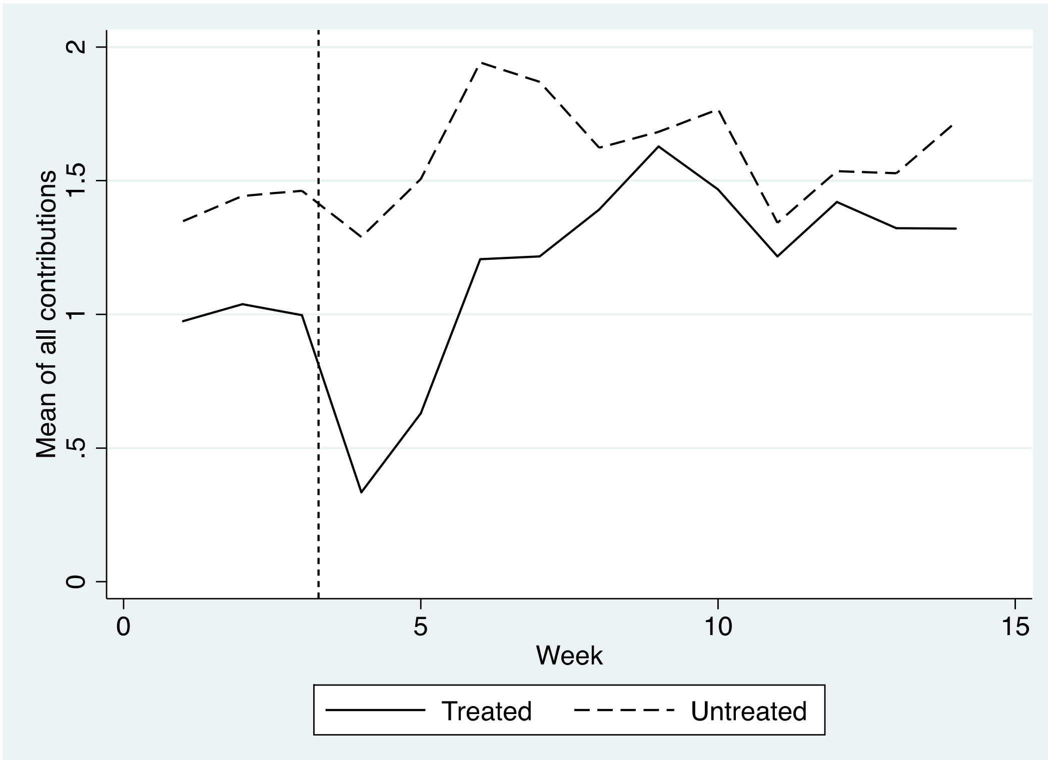

Figure 2: Open source productivity of workers in regions over time before and after treatment of lockdown
Note: Lockdown came in effect in Wuhan (Treated) on January 23rd and no lockdown was in place in Hong Kong and hence can be considered as untreated.

The regression results are shown in Table 3. For each dependent variable two models are developed here, first is the OLS regression model and the next is the fixed effects model. In model (1), the regressions uses only the independent variables and controls and we can see that the effect of treatment is significant ($p< .001$, $\beta= -.083$) on all contributions. In model (2), the analysis is repeated with fixed effects regression, the effect of treatment is significant ($p< .01$, $\beta= -.062$). In model (3), the same analysis with OLS regression is carried out for the entrepreneurship variable, we can see that the effect of treatment is partially significant ($p< .10$, $\beta= -.012$). In model (4), the analysis is repeated with fixed effects regression, the effect of treatment is significant ($p< .05$, $\beta= -.016$). In model (5) and (6) the analysis is repeated for personal learning dependent variable, the effect is significant in model (5) ($p< .001$, $\beta= -.053$) and model (6) ($p< .001$, $\beta= -.046$). In model (7) and (8) the analysis is repeated for project management dependent variable, the effect is significant in model (7) ($p< .001$, $\beta= -.039$) and model (8) ($p< .01$, $\beta= -.043$). Model (2) suggests that that lockdown treatment has decreased the all contributions of workers by 18.8% (see Kennedy (1981) for more details)(Kennedy 1981) and model (4) suggest 13% decrease in entrepreneurship, model (6) suggests 15.7% decrease in personal learning and model (8) suggest 15.5% decrease in project management activities.

Table 3: Results for effects of lockdown treatment on open source worker productivity

	(1) OLS	(2) FE	(3) OLS	(4) FE	(5) OLS	(6) FE	(7) OLS	(8) FE
Variables	All contributions		Entrepreneurship		Personal learning		Project management	
After	0.022** (0.007)	0.013 (0.010)	0.006* (0.002)	0.009* (0.004)	0.008* (0.004)	-0.004 (0.005)	0.014* (0.006)	0.017+ (0.009)
Location	0.037 (0.024)		-0.006 (0.007)		0.074*** (0.015)		-0.030 (0.018)	
$After_t$ * $Location_i$	-0.083*** (0.019)	-0.080*** (0.019)	-0.012+ (0.006)	-0.016* (0.007)	-0.053*** (0.013)	-0.046*** (0.014)	-0.039** (0.014)	-0.043** (0.015)
$Tenure_{it}$	0.097 (0.074)	0.067 (0.303)	0.036+ (0.020)	-0.029 (0.105)	0.005 (0.043)	0.020 (0.162)	0.116* (0.059)	-0.107 (0.262)
$(Tenure)_{it}$ ^2	0.000 (0.011)	0.022 (0.044)	-0.003 (0.003)	-0.000 (0.015)	0.005 (0.006)	0.026 (0.024)	-0.007 (0.008)	0.020 (0.038)
COVID-19 $infections_{it}$	0.010*** (0.003)	0.010** (0.003)	0.002* (0.001)	0.003* (0.001)	0.006*** (0.002)	0.005** (0.002)	0.006** (0.002)	0.007** (0.002)
Fixed Effects	No	Yes	No	Yes	No	Yes	No	Yes
Observations	43433	43433	43433	43433	43433	43433	43433	43433
Log likelihood	-44941	-21776	-604	14194	-17045	155	-38280	-15560
R squared	0.017	0.002	0.004	0.001	0.015	0.002	0.010	0.001

Notes: Standard errors in parentheses: + p<0.10, * p<0.05, ** p<0.01, *** p<0.001 robust standard errors clustered on individuals are reported here.

To further understand the time varying effects of the disruption we chose to study two weeks post treatment sequentially i.e., We compared the weeks (1-3) to week 5 and week 6 in first model (for t+1) and week 6 and 7 in second model for (t+2) (see Table 4 below) and we continued analyzing till the end of the lockdown period. From table 4, we can see that the effect of treatment was significant on all the dependent variables after t+1, though the effect vanished in the t+2 and it was significant and negative in t+3 weeks. Surprisingly, it took almost t+7 weeks to the workers to be productive again, we can see that entrepreneurship and project management activities becoming significant in the later half or close to the end of lockdown. This suggests that disruption effects only reverse after a very long time.

Table 4: Time varying effects of disruption

	All contributions	Entrepreneurship	Personal learning	Project management
t+1	-0.090***(0.022)	-0.028***(0.008)	-0.040**(0.015)	-0.062***(0.017)
t+2	-0.055(0.039)	-0.007(0.018)	-0.020(0.031)	-0.036(0.027)
t+3	-0.082*(0.033)	0.003(0.014)	-0.049+(0.026)	-0.034(0.024)
t+4	-0.189(0.148)	-0.008(0.042)	-0.151(0.112)	-0.052(0.092)
t+5	-0.208(0.192)	0.017(0.032)	-0.257(0.184)	0.057(0.056)
t+6	-0.260(0.191)	0.006(0.013)	-0.259(0.186)	0.021(0.025)
t+7	0.017(0.020)	0.016+(0.009	0.000(0.016)	0.030+(0.018)
t+8	0.021(0.022)	0.011(0.009)	0.000(0.017)	0.035+(0.019)
t+9	0.026(0.023)	0.013+(0.008)	-0.007(0.017)	0.041*(0.020)

Notes: All the models are fixed effects and the coefficients and standard errors of the interaction (after * location) are shown (standard errors are clustered on individuals). All the other variables (tenure, tenure square, COVID-19 infections) are included as controls.

Overall, the results suggest that the disruption due to COVID-19 pandemic has reduced the productivity of open source workers and the effect was found to be smaller in entrepreneurial activities than project management activities.

DISCUSSION AND CONCLUSIONS

The results of this study show that disruptions have a negative impact on the productivity of open source workers, though the effect was found to be lesser on entrepreneurial activities. Prior studies in online communities have found that workers self-organize during the crisis situations and our results show that in the early stages the negative effects are stronger (Nan and Lu 2014). Future research can study the aspect of self-organizing in open source development and understand the process by which workers become resilient. This is left for future study. Second, we find that the effect of disruption on project management activities and personal learning is much stronger (~16%) and hence researchers in open source need to unpack how to motivate workers during disruption and the possible mechanisms for enhancing community engagement (Daniel et al. 2013) and improving coordination (Lindberg et al. 2016).

The effect is consistently negative and well identified across various models by taking into account endogeneity issues (Nagle 2019). As this study only takes into open source platform activity, naturally it does not account for any other productivity related tasks that are not recorded by the platform and hence the true effect of lockdown on contributions of open source worker can be possibly much greater. Because of the sample construction and data constraints, there could be some open source workers who are not accounted for in the study. It is quite possible, that the users who are not accounted for in the study might show some differences in productivity levels, which are left for future study. Future research can explore how the productivity of the worker could differ across different user profiles (for example full-time vs. part-time worker).

An additional limitation relates to understanding on how productivity might vary across workers from different geographies as this study only considers two different locations. For example, future research can explore and conduct experiments in understanding how the productivity can gets affected by various lockdown policies and regional settings. Although, endogeneity is a concern in studies on productivity, this study uses a difference-in-difference technique and thus provides a causal explanation for decrease in productivity (Choudhury et al. 2020; Zhang and Zhu 2011).

The study findings have implications for researchers, policy makers, managers and organizations working in open source platforms. Disruptions are extremely common these types in open source platforms, hence platform owners need to understand what is the true effect of disruption. The answers to these questions can solve the puzzle of how platform owners can interact with the members of open source community by understanding the productivity losses. Even though, there was around 19% overall loss of contributions, the effective loss in starting new repositories (was only around 13%) suggesting that entrepreneurial activities may be less affected by disruption. Furthermore, the study suggest disruption effects on open source platforms due to the COVID-19 pandemic are much longer than expected. Hence, platform owners need to be accommodative and reassign tasks to less disrupted workers in case of crisis situations for increasing team productivity.

ACKNOWLEDGEMENTS

This work was supported with the financial support of the Science Foundation Ireland grant 13/RC/2094 and co-funded under the European Regional Development Fund through the Southern & Eastern Regional Operational Programme to Lero - the Science Foundation Ireland Research Centre for Software (www.lero.ie)

REFERENCES

Belenzon, S., and Schankerman, M. 2015. "Motivation and Sorting of Human Capital in Open Innovation," *Strategic Management Journal* (36:6), pp. 795-820.

Bertrand, M., Duflo, E., and Mullainathan, S. 2004. "How Much Should We Trust Differences-in-Differences Estimates?," *The Quarterly journal of economics* (119:1), pp. 249-275.
Choudhury, P., Koo, W. W., and Li, X. 2020. "Working (from Home) During a Crisis: Online Social Contributions by Workers During the Coronavirus Shock,").
Crowston, K., Annabi, H., and Howison, J. 2003. "Defining Open Source Software Project Success," *ICIS 2003 Proceedings*), p. 28.
Crowston, K., and Howison, J. 2005. "The Social Structure of Free and Open Source Software Development," *First Monday* (10:2), pp. 1–100.
Crowston, K., Howison, J., and Annabi, H. 2006. "Information Systems Success in Free and Open Source Software Development: Theory and Measures," *Software Process: Improvement and Practice* (11:2), pp. 123-148.
Daniel, S., Agarwal, R., and Stewart, K. J. 2013. "The Effects of Diversity in Global, Distributed Collectives: A Study of Open Source Project Success," *Information Systems Research* (24:2), pp. 312-333.
Dong, Y., Chung, M., Zhou, C., and Venkataraman, S. 2018. "Banking on “Mobile Money”: The Implications of Mobile Money Services on the Value Chain," *Manufacturing & Service Operations Management*).
Faraj, S., Jarvenpaa, S. L., and Majchrzak, A. 2011. "Knowledge Collaboration in Online Communities," *Organization science* (22:5), pp. 1224-1239.
Faraj, S., von Krogh, G., Monteiro, E., and Lakhani, K. R. 2016. "Special Section Introduction—Online Community as Space for Knowledge Flows," *Information systems research* (27:4), pp. 668-684.
Fisher, G. 2019. "Online Communities and Firm Advantages," *Academy of Management Review* (44:2), pp. 279-298.
Foss, N. J., Frederiksen, L., and Rullani, F. 2016. "Problem-Formulation and Problem-Solving in Self-Organized Communities: How Modes of Communication Shape Project Behaviors in the Free Open-Source Software Community," *Strategic management journal* (37:13), pp. 2589-2610.
Gousios, G. 2013. "The Ghtorent Dataset and Tool Suite," *Proceedings of the 10th Working Conference on Mining Software Repositories*: IEEE Press, pp. 233-236.
Horton, J. J., and Tambe, P. 2015. "Labor Economists Get Their Microscope: Big Data and Labor Market Analysis," *Big data* (3:3), pp. 130-137.
Kam Ng, M. 2008. "From Government to Governance? Politics of Planning in the First Decade of the Hong Kong Special Administrative Region," *Planning Theory & Practice* (9:2), pp. 165-185.
Kennedy, P. E. 1981. "Estimation with Correctly Interpreted Dummy Variables in Semilogarithmic Equations [the Interpretation of Dummy Variables in Semilogarithmic Equations]," *American Economic Review* (71:4), pp. 801-801.
Lindberg, A., Berente, N., Gaskin, J., and Lyytinen, K. 2016. "Coordinating Interdependencies in Online Communities: A Study of an Open Source Software Project," *Information Systems Research* (27:4), pp. 751-772.
Liu, Y., Gayle, A. A., Wilder-Smith, A., and Rocklöv, J. 2020. "The Reproductive Number of Covid-19 Is Higher Compared to Sars Coronavirus," *Journal of travel medicine*).
Medappa, P. K., and Srivastava, S. C. 2019. "Does Superposition Influence the Success of Floss Projects? An Examination of Open-Source Software Development by Organizations and Individuals," *Information Systems Research* (30:3), pp. 764-786.
Mora, R., and Reggio, I. 2012. "Treatment Effect Identification Using Alternative Parallel Assumptions,").
Morgeson, F. P., and DeRue, D. S. 2006. "Event Criticality, Urgency, and Duration: Understanding How Events Disrupt Teams and Influence Team Leader Intervention," *The Leadership Quarterly* (17:3), pp. 271-287.
Nagle, F. 2019. "Open Source Software and Firm Productivity," *Management Science* (65:3), pp. 1191-1215.
Nan, N., and Lu, Y. 2014. "Harnessing the Power of Self-Organization in an Online Community During Organizational Crisis," *Mis Quarterly* (38:4), pp. 1135-1158.
Puranam, P., Alexy, O., and Reitzig, M. 2014. "What's “New” About New Forms of Organizing?," *Academy of Management Review* (39:2), pp. 162-180.
Stewart, K. J., and Gosain, S. 2006. "The Impact of Ideology on Effectiveness in Open Source Software Development Teams," *Mis Quarterly*), pp. 291-314.
Stol, K.-J., and Fitzgerald, B. 2014. "Inner Source--Adopting Open Source Development Practices in Organizations: A Tutorial," *IEEE Software* (32:4), pp. 60-67.
Stoverink, A. C., Kirkman, B. L., Mistry, S., and Rosen, B. 2020. "Bouncing Back Together: Toward a Theoretical Model of Work Team Resilience," *Academy of Management Review* (45:2), pp. 395-422.
Sykes, E. R. 2011. "Interruptions in the Workplace: A Case Study to Reduce Their Effects," *International Journal of Information Management* (31:4), pp. 385-394.

Von Hippel, E., and Von Krogh, G. 2003. "Open Source Software and the" Private-Collective" Innovation Model: Issues for Organization Science," *Organization Science*), pp. 209-223.

Wooldridge, J. 2007. "What's New in Econometrics? Lecture 10 Difference-in-Differences Estimation," *NBER Summer Institute, available at: www. nber. org/WNE/Slides7–31–07/slides_10_diffindiffs. pdf, accessed April* (9), p. 2011.

Wu, J. T., Leung, K., Bushman, M., Kishore, N., Niehus, R., de Salazar, P. M., Cowling, B. J., Lipsitch, M., and Leung, G. M. 2020. "Estimating Clinical Severity of Covid-19 from the Transmission Dynamics in Wuhan, China," *Nature Medicine* (26:4), pp. 506-510.

Zhang, X. M., and Zhu, F. 2011. "Group Size and Incentives to Contribute: A Natural Experiment at Chinese Wikipedia," *American Economic Review* (101:4), pp. 1601-1615.