The Impacts of Lockdown on Open Source Software Contributions during the COVID-19 Outbreak

Jin Hu	Daning Hu	Xuan Yang	Michael Chau
The University of Hong Kong	Southern University of Science and Technology	University of Zurich	The University of Hong Kong
jinhu@connect.hku.hk	hdaning@gmail.com	xuanyang06@gmail.com	mchau@business.hku.hk

Abstract

We leverage the lockdown of Wuhan, China in January 2020 in response to COVID-19 as a natural experiment to study its impacts on individuals' contributions to open source software (OSS) on GitHub – the world's largest OSS platform. We find that Wuhan developers' contributions decreased by 10.2% relative to those in Hong Kong, Macau, and Taiwan (HMT) regions in the five weeks after the lockdown. Moreover, the contributions of Wuhan developers who interacted more with local developers on GitHub were reduced more after the lockdown. We conjecture that the lack of face-to-face (F2F) collaboration for Wuhan developers is the main driver of their reduced contributions, providing important insights for OSS platforms and stakeholders.

1. Introduction

In response to the threat of coronavirus disease 2019 (COVID-19), governments around the world have implemented lockdowns to reduce the risk of transmission [1]. The impacts of lockdowns on various individual behaviors have been studied, such as human mobility [2-4], work productivity [5-7], consumption [8, 9], Internet searching [10], and domestic violence [11, 12]. However, we still lack empirical evidence on the causal impact of the lockdown in response to COVID-19 on individuals' contributions to open source software (OSS), even though "there seemed to be a strong correlation between work from home transitions related lockdowns and measurable declines in developer activity" [13]. Moreover, research on working from home has shown mixed findings on developers' productivity during the COVID-19 pandemic [13-15], which also implies that the impact of the lockdown on developers' OSS contributions is not straightforward and needs to be explored. As digital public goods such as OSS contributions are especially needed in our societies in times of crisis like the COVID-19 pandemic

[16, 17], investigating the impact of lockdown on OSS contributions becomes important and meaningful.

The lockdown of Wuhan, China as a result of the COVID-19 outbreak provides us a unique natural experiment opportunity to study the impact of the lockdown on individuals' contributions to OSS on GitHub — the world's largest OSS platform [18]. Starting from January 23rd, 2020, the Chinese government implemented an unprecedented city-wide lockdown in Wuhan to prevent the spread of the COVID-19 virus. During the lockdown, all the Wuhan residents were required to stay at home. On the other hand, unlike most Chinese mainland cities that quickly followed Wuhan to implement similar social distancing (SD) measures, Hong Kong, Macau, and Taiwan (HMT) regions did not implement any lockdowns or other SD measures until March 2020. Moreover, compared with developers in other parts of the world, HMT developers are much more similar to Wuhan developers as they are from the same ethnic group—Han Chinese [19] and all adopt Chinese characters [20]. Therefore, we can compare the contribution behaviors between 1,679 Wuhan developers and 1,679 HMT developers (matched using propensity scores) from December 19th, 2019 to February 27th, 2020, to identify the causal effects of the lockdown.

Unpacking the mechanisms behind OSS developers' contribution behavioral responses to the lockdown is a complex challenge. Various mechanisms may be at work at the same time. For instance, staying and working from home during the pandemic, individuals tend to allocate the saved commuting time toward more computer-use and prosocial behavior like contributing to OSS [8, 16, 21]. Individuals may repurpose GitHub as a COVID-19 archive to warn the public about the virus [17]. On the other hand, the lockdown might reduce individual OSS contributions due to the challenging transition to the new normal of COVID-19 (e.g., worse mental health [10, 22] and low internet bandwidth at home [15]) and their struggle for work-social and work-life balance [23]. Moreover, OSS

developers who used to meet other developers in person regularly might reduce their contributions due to the sudden lack of face-to-face (F2F) collaboration caused by the lockdown.

Our study discovered a significant negative effect of the lockdown on individuals' OSS contributions during the COVID-19 outbreak in China. Specifically, we find that the lockdown accounts for a 10.2% drop in Wuhan developers' contributions relative to HMT developers' contributions. We also find that Wuhan developers who interacted more with other Wuhan developers on GitHub before the lockdown contributed much less after the lockdown, suggesting that having F2F collaboration among local developers may play a critical role in their contribution motivations.

While our empirical analyses do not yet allow for disentangling the exact roles of various mechanisms discussed above, as a first step, we document a significant reduction in individuals' OSS contributions in response to the lockdown during the COVID-19 outbreak. We also highlight the importance of F2F collaboration in motivating such contributions. Our findings contribute to the emerging literature on lockdowns and OSS literature and provide insights for digital platforms and organizations that rely on voluntary contributions to adapt to the new normal of COVID-19.

2. Related research

Our study is closely related to two streams of literature: (1) lockdowns during the COVID-19 pandemic and (2) motives to contribute to OSS.

2.1. Lockdowns during the COVID-19 pandemic

Our study adds to the emerging literature on lockdowns during the COVID-19 pandemic by examining how the lockdown affects individuals' OSS contributions. Although various individual behaviors during the lockdown have been studied [4, 5, 10], how the lockdown affects individuals' contributions to OSS lacks empirical evidence.

Most literature on lockdowns focused on how human mobility changes after the lockdown to evaluate this policy's effectiveness. For instance, Ruiz-Euler et al. [2] showed a mobility gap, i.e., the decline in human mobility during the lockdown happened at different speeds for high- versus low-income groups in most cities, because the lockdown imposed low-income groups with a tough choice between income and health. Durante et al. [3] found that mobility declined more in Italian provinces with higher civic capital, both before

and after the mandatory national lockdown. Fang et al. [4] found that the Wuhan lockdown substantially reduced inflows to and outflows from Wuhan and within-Wuhan movements. Overall, human mobility declines after the implementation of lockdown.

In addition to human mobility, other various individual behaviors during the lockdown have also been studied, such as work productivity [5-7], consumption [8, 9], Internet searching [10], and domestic violence [11, 12]. For example, Cui et al. [5] studied the impact of the lockdown in the United States on female and male academics' research productivity in social science. They found that female researchers' productivity dropped significantly relative to that of male researchers. This gender inequality in research productivity was also found among financial scholars [6]. Jaiswal and Arun [7] also showed reduced productivity levels of employees in India during the nationwide lockdown. Ghose et al. [8] found that consumers were more willing to share their mobile location data after the lockdown in the United States to help the public official flag hot spots of COVID-19. Coibion et al. [9] suggested a high correlation between the decline in consumer spending and a self-reported measure of whether the consumer was under a lockdown. Brodeur et al. [10] found that COVID-19 and the associated lockdowns implemented in Europe and America led to a substantial increase in the search intensity for negative words such as boredom and loneliness. Moreover, many studies found large increases in domestic violence during the lockdown [11, 12]. However, relatively little literature has focused on how the lockdown affects individuals' OSS contribution behaviors.

2.2. Motives to contribute to OSS

Our study also adds to the literature on motives for contributing to OSS by highlighting F2F collaboration as an important motivation, which has been relatively neglected in existing studies.

According to previous literature, incentives for OSS contributions can be categorized into intrinsic motivations, extrinsic motivations, and social effects. Intrinsic motivations are usually about contributors' personal needs such as altruism and joy [24, 25], while extrinsic motivations are usually related to external utility such as rewards and career advancement opportunities [26, 27]. Social effects refer to the phenomenon that contributors are influenced by others and receive social benefits from contributing [20]. For example, individuals' OSS contributions are encouraged by the attention they received [18] and collaboration with other contributors [28-30].

Online collaboration is a widely-discussed motive to contribute to OSS, as OSS developers usually collaborate to produce a software product on the OSS platform [26]. For instance, Xu et al. [28] found that OSS contributors build good interpersonal relationships and share ideology with other contributors through collaboration, which motivates them to contribute. Daniel and Stewart [29] suggested that collaboration helps contributors delegate tasks and avoid duplication of work in OSS projects, thus improving their efficiency and contribution levels. These studies focused on online collaboration probably because it makes contributors transcend the boundaries of time, distance, and organizations to collaborate [31].

However, only a few studies have focused on F2F collaboration, even though online collaboration alone cannot provide physical and emotional cues that help promote contributors' innovation and motivation [32]. Crowston et al. [30] suggested that F2F collaboration could greatly enhance OSS contributors' cooperation efficiency and increase their contributions. OSS literature also supported that F2F collaboration can "facilitate the transfer of tacit knowledge, which is more difficult in the context of computer-mediated relationships" [33] and improve team performance [34]. But overall, F2F collaboration has been relatively less discussed as a motive for contributing to OSS.

3. Empirical analyses

In this section, we first illustrate the setting and design of our natural experiment. Then we describe our data and construct the treatment and control groups for our natural experiment research design using propensity score matching (PSM). Moreover, we introduce our difference-in-differences (DID) regression to examine the causal impact of the lockdown on individuals' OSS contributions and present the empirical results. At last, we investigate whether the lack of F2F collaboration is the mechanism behind their contribution behavioral responses to the lockdown.

3.1. Experimental setting and design

COVID-19 has become one of the worst global pandemics in decades [4]. The virus was first discovered in Wuhan, China in early December of 2019, and has since rapidly spread around the world mainly through F2F interpersonal contacts [35]. From January 23rd to April 8th, 2020, Chinese government has implemented an unprecedented city-wide lockdown in Wuhan. During this lockdown, all types of transport were banned. All of the 7,148 residential communities in Wuhan were closed off, and residents were not allowed to leave Wuhan [4].

During this period, major cities in mainland China have implemented different levels of mandatory SD measures. For instance, public transport was temporarily shut down, and residential communities have adopted close-off management in Zhengzhou, Hangzhou, Harbin, Fuzhou, Zhumadian, and Ningbo staring from February 4th, 2020. In late February, some forms of SD measures were adopted and strictly enforced in most parts of mainland China.

In contrast, while tightening the entry restrictions and implementing quarantine rules, HMT regions did not implement any lockdowns or other SD measures until March 2020. Starting from March 29th, 2020, Hong Kong government banned indoor and outdoor public gatherings of more than four people. In Macau, although the local government has taken ad-hoc measures such as closing casinos and public parks have been implemented, no society-wide mandatory SD measures have been implemented during the COVID-19 outbreak. Therefore, from January 23rd, 2020 to March 30th, 2020, while developers in Wuhan had to stay at home, those in HMT regions could move freely and meet others in person.

As the world's largest OSS platform, GitHub provides an ideal empirical setting to study the impact of the lockdown on individuals' OSS contributions. First, all the contribution history of all developers and projects on GitHub can be accessed through API and GH Archive database. This allows us to track and measure the contributions from developers over time accurately. Second, the lockdown of Wuhan offers us a unique natural experiment to examine the causal impact of the lockdown. The lockdown is exogenous. There was no warning beforehand to anyone to avoid public panic and people leaving Wuhan to spread the virus further. Moreover, as the first city that was hugely affected by COVID-19 and implemented the lockdown, most people in Wuhan did not expect such strict policy on movement during the biggest holiday – the Chinese New Year. At the same time, individuals in HMT regions with no SD measures in place then have not realized and predicted that COVID-19 would also significantly affect their daily lives.

This natural experiment setting allows us to examine the changes in the contribution levels of Wuhan developers (the treatment group) and compare them with those of HMT developers (the control group) before and after five weeks of the Wuhan lockdown. We chose HMT developers as the control group for the following reasons. First, as mentioned before, most major cities in mainland China followed Wuhan very quickly to implement strict SD measures, while HMT regions did not do so until late March. Second, compared with developers in other parts of the world, HMT developers are much more similar to Wuhan

developers as they are from the same ethnic group—Han Chinese [19] and all adopt Chinese characters [20]. Therefore, by comparing the changes in the contribution levels of Wuhan developers and HMT developers before and after the Wuhan lockdown, we can study how Wuhan developers' contributions change as a result of the lockdown.

We chose the ten weeks surrounding the Wuhan lockdown (i.e., between December 19th, 2019 and February 27th, 2020) to construct our sample mainly for two reasons. First, the time window should not be too short to observe possible changes in developers' contributions. Second, as COVID-19 further spread to other parts of the world, including HMT regions, their developers may start to consciously avoid meeting others in person for the rising probability to be infected with the virus, even when their local authorities have not implemented mandatory SD measures. This would make HMT developers become less ideal control subjects in a natural experiment. Therefore, we set end of the time window as Feb 27th, 2020 as COVID-19 cases in HMT regions only started to increase in March substantially.

3.2. Data

Our dataset consists of two types of data: GitHub data and COVID-19 related data. We obtained historical GitHub data through the GitHub API and GH Archive database. The latter archives public GitHub OSS development history from February 2011 and has been widely used in recent OSS studies [18, 36]. We acquired COVID-19 related data mainly from the local health authorities such as the Wuhan Municipal Health Commission or Department of Health in Hong Kong and mainstream media.

We used the "search-by-location" function of the GitHub API to extract developers whose profile location is one of the four regions – Wuhan, Hong Kong, Macau, and Taiwan. We then crawled the profile information of developers who joined GitHub before December 19th, 2019. This procedure yielded 1,689 Wuhan developers and 4,434 HMT developers. We gathered data from the GH Archive database about the contributions and other GitHub activities of these 6,123 developers from February 2011 (when GH Archive started to archive GitHub data) to February 27th, 2020 (the end of the time window).

3.3. Matching

To correct potential sample selection bias due to any observable differences between a Wuhan developer and an HMT developer, we applied the DID technique in combination with PSM [37, 38]. We matched each Wuhan developer with an HMT developer who is similar in terms of the following observed characteristics on GitHub following previous studies [18, 20, 26]: the number of *weeks* since the developer joined GitHub until the day of Wuhan lockdown, whether the developer is a *student* or an *employee* based on her profile, whether the developer reports her *contact* information in her profile, the number of public *repositories* that the developer created before the lockdown, the number of *commits* that the developer contributed on GitHub before the lockdown, the number of *stars/issues/comments* that the developer *received* for her repositories before the lockdown, and the number of *stars/issues/comments* that the developer *sent out* before the lockdown.

Table 1 reports the results of *t*-tests for the variables used in PSM, both before (Panel A) and after (Panel B) the matching. Overall, the PSM reduced the observable differences in trends and balanced the group sizes, yielding 1,679 treated Wuhan developers and 1,679 matched control HMT developers.

3.4. Impact of the lockdown on contributions

We estimate the impact of the lockdown on individuals' OSS contributions using the following DID specification:

$$CONTRIBUTION_{it} = \alpha + \beta AFTER_t \\ \times LOCKDOWN_i \\ + \gamma CV_{it} + \mu_i + \theta_t + \epsilon_{it} \qquad (1)$$

where i indexes the developer and t indexes the week. The dependent variable, $CONTRIBUTION_{it}$, is the weekly contributions of each developer on GitHub. We added one to the weekly number of commits a developer contributed on GitHub and then took a logarithm to measure her weekly contributions following previous literature [18, 20, 39]. A commit is a change to an OSS repository, such as adding, modifying, and deleting codes. $AFTER_t$ is a dummy that equals one if the period is after January 23rd, 2020, and zero otherwise. $LOCKDOWN_i$ is a dummy that equals one for the developers in Wuhan and zero otherwise. CV_{it} contains the following control variables that might influence individual contribution behavior according to previous literature [18, 20, 26]: the number of public repositories created by the developer ($REPO_{it}$), the number of weeks since the developer joined GitHub ($TENURE_{it}$), the number of stars that the developer received for her repositories ($STARR_{it}$), the number of stars that the developer sent out ($STARS_{it}$), the number of issues that the developer received for her repositories ($ISSUER_{it}$), the number of issues that the developer sent out ($ISSUES_{it}$), the number of comments that the developer received for her repositories ($COMMENTR_{it}$), the

number of comments that the developer sent out ($COMMENTS_{it}$), and the number of new confirmed COVID-19 cases in the developer' region ($CASE_{it}$). μ_i is the individual fixed effect. θ_t is the time fixed effect. ϵ_{it} is the error term.

Table 1. *T*-tests in the overall and PSM-matched samples

Variable	Mean		*t*-test
	Wuhan	HMT	Difference
Panel A: Overall sample			
weeks	174.761	224.249	-49.488***
student	0.305	0.162	0.143***
employee	0.232	0.290	-0.058***
contact	0.721	0.690	0.031**
repositories	21.780	26.074	-4.294***
commits	709.337	1489.132	-779.795**
stars received	126.292	77.595	48.697
issues received	6.959	9.129	-2.169
comments received	12.684	24.629	-11.945**
stars sent out	104.883	107.217	-2.334
issues sent out	8.740	12.421	-3.681*
comments sent out	23.530	47.056	-23.526**
Panel B: PSM-matched sample			
weeks	176.136	170.362	5.774*
student	0.299	0.273	0.026
employee	0.233	0.247	-0.014
contact	0.718	0.718	0.000
repositories	21.814	22.340	-0.526
commits	713.119	693.205	19.914
stars received	81.420	39.272	42.148***
issues received	6.803	5.081	1.722
comments received	12.361	9.983	2.378
stars sent out	105.431	101.066	4.365
issues sent out	8.756	7.286	1.469
comments sent out	23.537	20.872	2.665

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

Table 2 reports the estimated results of Equation (1). The estimated coefficient of $AFTER_t \times LOCKDOWN_i$ suggests that Wuhan developers' contributions decreased by 10.2% ($= e^{-0.108} - 1$) over the five-week period after the lockdown. This indicates that the lockdown has a negative effect on developers' OSS contributions.

Table 2. Regression results for the DID model

Variable	Dependent variable: $CONTRIBUTION_{it}$
$AFTER_t \times LOCKDOWN_i$	-0.108*** (0.028)
$REPO_{it}$	0.396*** (0.051)
$TENURE_{it}$	-0.007 (0.037)
$STARR_{it}$	0.018*** (0.006)
$STARS_{it}$	0.011** (0.005)
$ISSUER_{it}$	0.009 (0.025)
$ISSUES_{it}$	0.030 (0.021)
$COMMENTR_{it}$	0.031*** (0.010)
$COMMENTS_{it}$	0.048*** (0.010)
$CASE_{it}$	-0.000 (0.000)
Constant	1.790 (6.204)
Individual FE	Yes
Time FE	Yes
Observations	33580
R-squared	0.114

Robust standard errors in brackets.
* p<0.1, ** p<0.05, *** p<0.01.

3.5. Lack of F2F collaboration

Wuhan Developers who may meet each other in person regularly can be significantly affected by the sudden lack of F2F collaboration caused by the lockdown. Previous research suggests that F2F collaboration enhances OSS contributors' efficiency and motivates them to contribute [30, 33, 34].

Our study assumes that Wuhan developers who had more online interactions among themselves (i.e., made more comments to other Wuhan developers on GitHub) were more likely to meet F2F before the lockdown. Therefore, the contributions of these developers are expected to be reduced more after the lockdown since they relied more on F2F collaboration and would be affected more by the lockdown.

To examine whether the lack of F2F collaboration is the mechanism by which Wuhan developers reduced contributions after the lockdown, we consider the

following difference-in-difference-in-differences (DDD) specifications:

$$CONTRIBUTION_{it} = \alpha + \beta_1 AFTER_t \times LOCKDOWN_i + \beta_2 AFTER_t \times LOCCOMS_i + \beta_3 AFTER_t \times LOCKDOWN_i \times LOCCOMS_i + \gamma CV_{it} + \mu_i + \theta_t + \epsilon_{it} \quad (2)$$

where $LOCCOMS_i$ is the number of comments that the developer made to other GitHub developers in the same region before the lockdown.

$$CONTRIBUTION_{it} = \alpha + \beta_1 AFTER_t \times LOCKDOWN_i + \beta_2 AFTER_t \times COMS_i + \beta_3 AFTER_t \times LOCKDOWN_i \times COMS_i + \gamma CV_{it} + \mu_i + \theta_t + \epsilon_{it} \quad (3)$$

where $COMS_i$ is the number of comments that the developer made to all GitHub developers before the lockdown.

Table 3 reports the estimated results of Equations (2) and (3) in Column (1) and (2), respectively. The estimated coefficient of $AFTER_t \times LOCKDOWN_i \times LOCCOMS_i$ in Column (1) is significant and negative, which indicates that Wuhan developers who had more online interactions with other Wuhan developers contributed much less after the lockdown. We suggest this reduction is very likely caused by the lack of F2F collaboration among these local Wuhan developers. It is consistent with previous literature which supports that F2F collaboration can enhance OSS contributors' efficiency and increase their contributions [30, 33, 34]. One alternative explanation is that developers who had more online interactions with others are more social and would be affected more by the general well-being of others in their lives. For instance, these developers might be more worried or stressed about the pandemic and thus had less energy to contribute to OSS [40]. However, this alternative was ruled out since the estimated coefficient of $AFTER_t \times LOCKDOWN_i \times COMS_i$ in Column (2) is insignificant, showing that the contributions of Wuhan developers who had more online interactions with other developers (inside and outside Wuhan) were not affected more (or less) by the lockdown.

4. Robustness checks

Table 3. Regression results for the DDD models

Variable	(1)	(2)
	Dependent variable: $CONTRIBUTION_{it}$	
$AFTER_t \times LOCKDOWN_i$	-0.106*** (0.028)	-0.109*** (0.028)
$AFTER_t \times LOCCOMS_i$	0.000 (0.001)	
$AFTER_t \times LOCKDOWN_i \times LOCCOMS_i$	-0.005** (0.003)	
$AFTER_t \times COMS_i$		-0.000 (0.000)
$AFTER_t \times LOCKDOWN_i \times COMS_i$		0.000 (0.000)
$REPO_{it}$	0.396*** (0.051)	0.396*** (0.051)
$TENURE_{it}$	-0.007 (0.037)	-0.008 (0.037)
$STARR_{it}$	0.018*** (0.006)	0.018*** (0.006)
$STARS_{it}$	0.011** (0.005)	0.011** (0.005)
$ISSUER_{it}$	0.010 (0.025)	0.010 (0.026)
$ISSUES_{it}$	0.030 (0.021)	0.030 (0.021)
$COMMENTR_{it}$	0.031*** (0.010)	0.031*** (0.010)
$COMMENTS_{it}$	0.048*** (0.010)	0.048*** (0.010)
$CASE_{it}$	-0.000 (0.000)	-0.000 (0.000)
Constant	1.802 (6.206)	1.878 (6.202)
Individual FE	Yes	Yes
Time FE	Yes	Yes
Observations	33580	33580
R-squared	0.114	0.114

Robust standard errors in brackets.
* p<0.1, ** p<0.05, *** p<0.01.

In this section, we report the results of several robustness tests. Specifically, we check the parallel trends assumption, conduct falsification tests, and replicate the DID analysis using an alternative sample.

4.1. Parallel trends

The key identification assumption for the DID estimation is the parallel trends assumption: before the Wuhan lockdown, Wuhan developers' and HMT

developers' contributions on GitHub would follow the same time trend.

Figure 1 plots the weekly average contributions (per developer) made by Wuhan developers (treatment group; blue) and matched HMT developers (control group; red) five weeks before and after the day of Wuhan lockdown. To measure a developer's weekly contributions, we added one to the weekly number of commits she contributed on GitHub and then took a logarithm, consistent with the measure in Equation (1). The green vertical line represents the day of Wuhan lockdown. We can observe that, before the lockdown began, both lines are very similar to each other, which fulfills the parallel trends assumption. On the other hand, there is consistently a large gap between the red and blue lines in each of the five weeks after the lockdown, suggesting Wuhan developers, on average, made much fewer contributions than HMT developers during that time.

Figure 1. Time trends of weekly average contributions on GitHub

To further validate the parallel trends assumption, we also adopt an event-study approach by fitting the following equation [22]:

$$CONTRIBUTION_{it} = \alpha$$
$$+ \sum_{k=-5, k \neq -1, k \neq 0}^{5} \beta_k$$
$$\times (WEEK_{t,k}$$
$$\times LOCKDOWN_i)$$
$$+ \gamma CV_{it} + \mu_i + \theta_t + \epsilon_{it} \qquad (4)$$

where $WEEK_{t,k}$ is a dummy that equals one if week t corresponds to k, and zero otherwise. We did not construct the week $k = 0$ in our sample but use the day of Wuhan lockdown to separate the pre-treatment and post-treatment period. $k = -1$ indicates the week just before the day of Wuhan lockdown, so it is dropped from the equation as the reference week. Intuitively, β_k captures the difference in contributions between Wuhan developers and HMT developers in each week relative

to $k = -1$. We expect the treatment group and control group to make similar contributions before the lockdown ($k < 0$) and to diverge after the lockdown ($k > 0$).

Figure 2 shows the estimated β_k in Equation (4). The green vertical line represents the day of Wuhan lockdown. The gray dotted lines for each coefficient show 95% confidence intervals. We can observe that the estimated β_{-5}, β_{-4}, β_{-3}, and β_{-2} are all almost zero, indicating no pre-treatment differences in the contribution trends between Wuhan developers and HMT developers, thus supporting the parallel trends assumption.
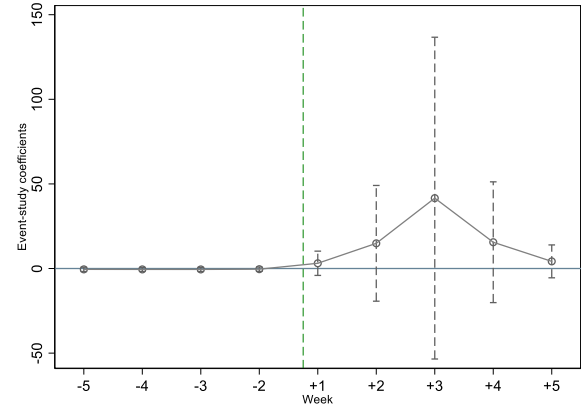

Figure 2. Event-study estimation

4.2. Falsification test

One concern is that our estimated effects are an artifact of seasonality, i.e., Wuhan developers contributed less than HMT developers during the Spring Festival holiday in any year. To rule out the seasonal effects, we repeat the same analysis specified in Equation (1) for the same period a lunar year ago. The control variable $CASE_{it}$ is not included this time because the COVID-19 did not break out at that time. If our results simply capture seasonality, we would be able to find a significant effect a lunar year ago.

Table 4 reports the falsification test results. The placebo-treated treatment effect is insignificant, implying that Wuhan developers did not reduce contributions significantly in the previous lunar year, thus ruling out the seasonal effects.

4.3. Alternative sample

We also replicate the DID analysis using an alternative sample based on coarsened exact matching (CEM). CEM matches based on ex-ante criteria, effectively minimizing heterogeneities between the treatment and control groups [41]. It has been used

extensively in recent work to mitigate endogeneity concerns [37, 38]. We perform the k-to-k CEM based on the same covariates used in PSM, yielding 1,634 Wuhan developers and 1,634 matched HMT developers. Table 5 shows the results of *t*-tests for the variables used in CEM after the matching. Compared with Panel A in Table 1, we can observe that the CEM narrows the differences of those variables between the treatment and control groups.

Table 4. Falsification test results

Variable	Dependent variable: $CONTRIBUTION_{it}$
$AFTER_t \times LOCKDOWN_i$	-0.007 (0.019)
$REPO_{it}$	0.185*** (0.043)
$TENURE_{it}$	0.052* (0.028)
$STARR_{it}$	0.011*** (0.003)
$STARS_{it}$	0.061*** (0.008)
$ISSUER_{it}$	-0.015 (0.020)
$ISSUES_{it}$	0.119*** (0.043)
$COMMENTR_{it}$	0.009 (0.014)
$COMMENTS_{it}$	0.133*** (0.024)
Constant	-5.952* (3.374)
Individual FE	Yes
Time FE	Yes
Observations	32680
R-squared	0.163

Robust standard errors in brackets.
* p<0.1, ** p<0.05, *** p<0.01.

We then replicate the DID analysis specified in Equation (1) using the CEM-matched sample. We obtain consistent results, as shown in Table 6. The estimated coefficient of $AFTER_t \times LOCKDOWN_i$ is still significant and negative, supporting the negative impact of the lockdown on individuals' OSS contributions.

5. Discussion and conclusion

Leveraging the Wuhan lockdown during the COVID-19 outbreak as a natural experiment, we empirically examine the impacts of the lockdown on individuals' OSS contributions. We find that Wuhan developers have made 10.2% fewer contributions than HMT developers in the five weeks after the Wuhan lockdown. We also find that the contributions of Wuhan developers who interacted more with other Wuhan developers on GitHub were reduced more after the lockdown, suggesting that the reduction in Wuhan developers' contributions may be resulted from the lack of F2F collaboration caused by the lockdown.

Table 5. *T*-tests in the CEM-matched sample

Variable	Mean		*t*-test
	Wuhan	HMT	Difference
weeks	174.104	174.222	-0.118
student	0.299	0.299	0.000
employee	0.232	0.232	0.000
contact	0.714	0.714	0.000
repositories	19.009	17.931	1.078
commits	520.674	550.239	-29.565
stars received	42.473	33.634	8.839
issues received	4.369	3.606	0.763
comments received	6.486	7.882	-1.397
stars sent out	87.793	62.736	25.057***
issues sent out	6.284	5.274	1.010*
comments sent out	14.606	17.864	-3.258

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

Our study contributes to the emerging literature on lockdowns during the COVID-19 pandemic by investigating how the lockdown affects individuals' OSS contributions. Our findings show that the lockdown negatively affects individuals' OSS contributions. Our study also contributes to OSS literature by highlighting the importance of F2F collaboration (or lack of it) in motivating contributions.

Our findings have implications for digital platforms and organizations that rely on voluntary contributions and highlight the need for such platforms to design features to sustain contributions under the new normal of COVID-19. Relevant platforms and organizations may leverage advanced technology (e.g., adopting a facial micro-expression recognition system for online meetings) to remedy the limitations of online collaboration and facilitate efficient and effective collaboration among contributors.

Some limitations of our study generate directions and opportunities for future research. First, we have limited information about the GitHub developers. Information about developers' offline lives and detailed demographics such as gender is not available on GitHub. Our future work plans to use survey to collect

the sample Wuhan developers' information (e.g., F2F collaboration experience) around the lockdown to examine if Wuhan developers who made more comments to local developers on GitHub are more likely to meet F2F before the lockdown and to disentangle various mechanisms behind the reduction of the contributions. Second, our consideration of OSS contributions focuses on the amounts of commits. While it is a reasonable and widely used measure of OSS contributions [18, 39], it may not necessarily mean that the contributor is doing less (high-quality) work. The contributors might also focus more on other activities rather than contributing commits on GitHub. Future research may consider using content analysis to assess the nature of these commits and investigating other activities on GitHub.

Table 6. Regression results for the DID model using the CEM-matched sample

Variable	Dependent variable: $CONTRIBUTION_{it}$
$AFTER_t \times LOCKDOWN_i$	-0.088*** (0.027)
$REPO_{it}$	0.469*** (0.046)
$TENURE_{it}$	-0.021 (0.037)
$STARR_{it}$	0.019*** (0.007)
$STARS_{it}$	0.009* (0.005)
$ISSUER_{it}$	-0.020 (0.033)
$ISSUES_{it}$	0.089*** (0.018)
$COMMENTR_{it}$	0.021** (0.010)
$COMMENTS_{it}$	0.065*** (0.010)
$CASE_{it}$	-0.000 (0.000)
Constant	4.013 (6.209)
Individual FE	Yes
Time FE	Yes
Observations	32680
R-squared	0.129

Robust standard errors in brackets.
* $p<0.1$, ** $p<0.05$, *** $p<0.01$.

References

[1] J.A. Weill, M. Stigler, O. Deschenes, et al., "Social Distancing Responses to COVID-19 Emergency Declarations Strongly Differentiated by Income", Proceedings of the National Academy of Sciences of the United States of America, 117(33), pp. 19658-19660, 2020.

[2] A. Ruiz-Euler, F. Privitera, D. Giuffrida, et al., "Mobility Patterns and Income Distribution in Times of Crisis: US Urban Centers During the COVID-19 Pandemic", SSRN 3572324, 2020.

[3] R. Durante, L. Guiso, and G. Gulino, "Asocial Capital: Civic Culture and Social Distancing during COVID-19", Journal of Public Economics, 194, pp. 104342, 2021.

[4] H. Fang, L. Wang, and Y. Yang, "Human Mobility Restrictions and the Spread of the Novel Coronavirus (2019-nCoV) in China", Journal of Public Economics, 191, pp. 104272, 2020.

[5] R. Cui, H. Ding, and F. Zhu, "Gender Inequality in Research Productivity During the COVID-19 Pandemic", Manufacturing & Service Operations Management, 2021.

[6] B.M. Barber, W. Jiang, A. Morse, et al., "What Explains Differences in Finance Research Productivity During the Pandemic?", Journal of Finance, 2021.

[7] A. Jaiswal and C.J. Arun, "Unlocking the COVID-19 Lockdown: Work from Home and Its Impact on Employees", 2020.

[8] A. Ghose, B. Li, M. Macha, et al., "Trading Privacy for the Greater Social Good: How Did America React During COVID-19?", SSRN 3624069, 2020.

[9] O. Coibion, Y. Gorodnichenko, and M. Weber, "The Cost of the COVID-19 crisis: Lockdowns, Macroeconomic Expectations, and Consumer Spending", National Bureau of Economic Research Working Paper 27141, 2020.

[10] A. Brodeur, A. Clark, S. Fleche, et al., "COVID-19, Lockdowns and Well-Being: Evidence from Google Trends", Journal of Public Economics, 193, pp. 104346, 2021.

[11] E. Arenas-Arroyo, D. Fernandez-Kranz, and N. Nollenberger, "Intimate Partner Violence under Forced Cohabitation and Economic Stress: Evidence from the COVID-19 Pandemic", Journal of Public Economics, 194, pp. 104350, 2021.

[12] S. Ravindran and M. Shah, "Unintended Consequences of Lockdowns: COVID-19 and the Shadow Pandemic", National Bureau of Economic Research Working Paper 27562, 2020.

[13] V. Combs, "Productivity and WFH: Developers slow to bounce back worldwide as lockdown lifts", Available from: https://www.techrepublic.com/article/productivity-and-wfh-developers-slow-to-bounce-back-worldwide-as-lockdown-lifts/, 2020.

[14] L. Bao, T. Li, X. Xia, et al., "How does Working from Home Affect Developer Productivity? – A Case Study of Baidu During COVID-19 Pandemic", Proceedings of The 28th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering, ACM, New York, NY, USA, 2020.

[15] P.A.d.M.S. Neto, U.A. Mannan, E.S. de Almeida, et al., "A Deep Dive on the Impact of COVID-19 in Software Development", arXiv preprint arXiv:2008.07048, 2020.

[16] M. Kummer, O. Slivko, and X.M. Zhang, "Unemployment and Digital Public Goods Contribution", Information Systems Research, 31(3), pp. 801-819, 2020.

[17] Y.-L. Liu, "In China, GitHub Is a Free Speech Zone for Covid Information", Available from: https://www.wired.com/story/china-github-free-speech-covid-information/, 2020.

[18] M. Moqri, X. Mei, L. Qiu, et al., "Effect of "Following" on Contributions to Open Source Communities", Journal of Management Information Systems, 35, pp. 1188-1217, 2018.

[19] Wikipedia, "Han Chinese", Available from: https://en.wikipedia.org/wiki/Han_Chinese, 2021.

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

[21] N. Forsgren, "Octoverse Spotlight: An Analysis of Developer Productivity, Work Cadence, and Collaboration in the Early Days of COVID-19", Available from: https://github.blog/2020-05-06-octoverse-spotlight-an-analysis-of-developer-productivity-work-cadence-and-collaboration-in-the-early-days-of-covid-19/, 2020.

[22] T. Tanaka and S. Okamoto, "Increase in Suicide Following an Initial Decline during the COVID-19 Pandemic in Japan", Nature Human Behaviour, 5(2), pp. 229-238, 2021.

[23] P. Choudhury, W.W. Koo, and X. Li, "Working (From Home) During a Crisis: Online Social Contributions by Workers During the Coronavirus Shock", Harvard Business School Working Paper 20-096, 2020.

[24] J. Davidson, U. Mannan, R. Naik, et al., "Older Adults and Free/Open Source Software: A Diary Study of First-Time Contributors", Proceedings of the International Symposium on Open Collaboration, Association for Computing Machinery, New York, NY, United States, 2014.

[25] G. Hertel, S. Niedner, and S. Herrmann, "Motivation of Software Developers in Open Source Projects: an Internet-Based Survey of Contributors to the Linux Kernel", Research Policy, 32(7), pp. 1159-1177, 2003.

[26] Y. Fang and D. Neufeld, "Understanding Sustained Participation in Open Source Software Projects", Journal of Management Information Systems, 25, pp. 9-50, 2009.

[27] X. Yang, X. Li, D. Hu, et al., "Differential Impacts of Social Influence on Initial and Sustained Participation in Open Source Software Projects", Journal of the Association for Information Science and Technology, 72(9), pp. 1133-1147, 2021.

[28] B. Xu, D.R. Jones, and B. Shao, "Volunteers' Involvement in Online Community based Software Development", Information & Management, 46(3), pp. 151-158, 2009.

[29] S. Daniel and K. Stewart, "Open Source Project Success: Resource Access, Flow, and Integration", Journal of Strategic Information Systems, 25(3), pp. 159-176, 2016.

[30] K. Crowston, J. Howison, C. Masango, et al., "The Role of Face-to-Face Meetings in Technology-Supported Self-Organizing Distributed Teams", IEEE Transactions on Professional Communication, 50(3), pp. 185-203, 2007.

[31] B.J. Bergiel, E.B. Bergiel, and P.W. Balsmeier, "The Reality of Virtual Teams", Competition Forum, American Society for Competitiveness, 2006.

[32] R.K. Hart, "Constituting Relationships in Communication: An Interdisciplinary Approach to Understanding Peer Relationships in Geographically Dispersed Teams", Advances in Interdisciplinary Studies of Work Teams, 8, pp. 85-106, 2001.

[33] W. Stam, "When Does Community Participation Enhance the Performance of Open Source Software Companies?", Research Policy, 38(8), pp. 1288-1299, 2009.

[34] D. Hu, X. Li, and X. Zhang, "The Impacts of Geographic Dispersion on OSS Project Success: Face-to-face vs. Virtual Collaboration", Thirty Seventh International Conference on Information Systems, Dublin, Ireland, 2016.

[35] J.F.-W. Chan, S. Yuan, K.-H. Kok, et al., "A Familial Cluster of Pneumonia Associated with the 2019 Novel Coronavirus Indicating Person-to-Person Transmission: A Study of A Family Cluster", The Lancet, 395(10223), pp. 514-523, 2020.

[36] B. Negoita, G. Vial, M. Shaikh, et al., "Code Forking and Software Development Project Sustainability: Evidence from GitHub", Fortieth International Conference on Information Systems, Munich, Germany, 2019.

[37] J. Foerderer, "Interfirm Exchange and Innovation in Platform Ecosystems: Evidence from Apple's Worldwide Developers Conference", Management Science, 66(10), pp. 4772-4787, 2020.

[38] S. Ye, S. Viswanathan, and I.-H. Hann, "The Value of Reciprocity in Online Barter Markets: An Empirical Investigation", MIS Quarterly, 42(2), pp. 521-549, 2018.

[39] W. Chen, F. Jin, and L. Xue, "Flourish or Perish? The Impact of Technological Acquisitions on Contributions to Open-Source Software", Kelley School of Business Research Paper No. 18-55, 2018.

[40] C. Miller, D.G. Widder, C. Kästner, et al., "Why do People Give Up FLOSSing? A Study of Contributor Disengagement in Open Source", IFIP International Conference on Open Source Systems, Springer, 2019.

[41] S.M. Iacus, G. King, and G. Porro, "Causal Inference Without Balance Checking: Coarsened Exact Matching", Political Analysis, 20, pp. 1-24, 2012.