

Association for Information Systems

AIS Electronic Library (AISeL)

ICIS 2022 Proceedings

Sharing Economy, Platforms, and Crowds

Dec 12th, 12:00 AM

How Does Anonymizing Crowdsourced Users' Identity Affect Fact-checking on Social Media Platforms? A Regression Discontinuity Analysis

Guohou Shan

Temple University, guohoushan@temple.edu

Sunil Wattal

Temple University, swattal@temple.edu

Jason B. Thatcher

Temple University, jason.b.thatcher@gmail.com

Follow this and additional works at: <https://aisel.aisnet.org/icis2022>

Recommended Citation

Shan, Guohou; Wattal, Sunil; and Thatcher, Jason B., "How Does Anonymizing Crowdsourced Users' Identity Affect Fact-checking on Social Media Platforms? A Regression Discontinuity Analysis" (2022). *ICIS 2022 Proceedings*. 13.

https://aisel.aisnet.org/icis2022/sharing_econ/sharing_econ/13

This material is brought to you by the International Conference on Information Systems (ICIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in ICIS 2022 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

How Does Anonymizing Crowdsourced Users' Identity Affect Fact-checking on Social Media Platforms? A Regression Discontinuity Analysis

Completed Research Paper

Guohou Shan

Temple University
1801 N Broad St,
Philadelphia, PA 19122
guohou.shan@temple.edu

Sunil Wattal

Temple University
1801 N Broad St,
Philadelphia, PA 19122
swattal@temple.edu

Jason Bennett Thatcher

Temple University
1801 N Broad St,
Philadelphia, PA 19122
starlord@temple.edu

Abstract

The rapid spread of misinformation on social media platforms has affected many facets of society, including presidential elections, public health, the global economy, and human well-being. Crowdsourced fact-checking is an effective method to mitigate the spread of misinformation on social media. A key factor that affects user behavior on crowdsourcing platforms is users' anonymity or identity disclosure. Within the crowdsourced-based fact-checking context, it is also unknown whether and how identity anonymity affects the users' fact-checking contribution performance. Leveraging a natural experiment policy happening on Twitter, we adopt regression discontinuity design to investigate two research questions: Whether and how the identity anonymity affects the crowdsourced fact-checking quantity and quality; how the characteristics of the crowdsourced users moderate the main impact. We find that the identity anonymization policy may not increase fact-checking users' contribution quantity, but the fact-checking quality does increase. Our research has both theoretical and practical implications.

Keywords: crowdsourced fact-checking; identity anonymity; fact-checking quantity and quality; regression discontinuity design

Introduction

The rapid spread of misinformation on social media platforms (Acemoglu et al. 2021) has affected many facets of society, including presidential elections, public health, the global economy, and human well-being. For example, false claims and misinformation around voter fraud reached an unprecedented level on social media, potentially affecting the 2020 U.S. presidential election (Pennycook and Rand 2021). Also, 75% of U.S. adults seek health information online (mainly through social media channels), while only about 20% have the capability to distinguish misinformation (Rodgers and Massac 2020) from helpful advice. Furthermore, combating misinformation on social media can cost billions of dollars a year (Castillo 2019).

Crowdsourced fact-checking is used to mitigate the spread of misinformation on social media. Prior research (Allen et al. 2021a; Micallef et al. 2020; Pennycook and Rand 2019) has shown that crowdsourced

users (fact-checkers) can be similarly accurate and faster than professional fact-checkers in flagging misinformation, and can help reduce the spread of misinformation. One of the first social media platforms in the US to leverage crowdsourced-based fact-checking services was Twitter. On January 25, 2021, Twitter launched the Birdwatch program with the objective of using crowdsourced users to fight misinformation on its platform.¹

A key factor that affects user behavior on crowdsourcing platforms is users' anonymity or identity disclosure (Burtch et al., 2015). On the one hand, anonymizing users' identities can help protect user privacy, making users feel free to contribute more frequently. On the other hand, identity anonymization can also decrease the earned reputation and social recognition, which may decrease some crowd users' motivation to contribute more and higher quality content. Furthermore, literature has documented the contradicting impact of identity anonymity on crowdsourcing users' contributions. For example, studies have found identity anonymity can help users feel less concerned about being retaliated against in reporting disruptive events on social media, contributing more to reporting disruptive events (e.g., misinformation) (Lowry et al. 2013; Wong et al. 2021). However, anonymizing users' identity can inhibit their reputational gains, thus decreasing their contribution (e.g., in the crowdfunding context, anonymous donors are extremely rare) (Huang et al. 2017; Wang 2010).

Within the crowdsourced-based fact-checking context, it is also unknown whether and how identity anonymity affects the users' fact-checking contribution performance. Users on social media may have different reasons for flagging misinformation. For example, some Twitter Birdwatch program users think the recognition and reputation of being an expert in flagging a particular type of misinformation motivate them to fact-check and flag misinformation (Mahadevan and Mantas 2021). However, it is also possible that anonymizing the identity makes users worry less about being retaliated when fact-checking and flagging controversial misinformation (Mackey 2021).

Therefore, whether and how identity anonymization affects crowdsourced users' fact-checking behaviors is an open research question. For example, users may reduce their fact-checking behavior because they couldn't gain reputations via flagging misinformation. On the other hand, they may also increase their contribution because they would not be concerned about being retaliated against by reporting misinformation. From a theoretical standpoint, it is interesting to investigate which effect dominates in the context of crowdsourced fact-checking. Moreover, it is also interesting to understand whether the impact of identity anonymity of crowdsourced users can be moderated by the characteristics of the crowdsourced users (e.g., their contribution tenure). Different crowdsourced users may have different reasons and adopt different strategies in engaging with flagging misinformation. Therefore, the characteristics of the users may play a moderating role in the impact of identity anonymity. From a practical standpoint, firms such as social media platforms can design policies, guidelines, and targeted communications based on a better understanding of how anonymity impacts user behavior in crowdsourced fact-checking.

In this research, we leverage a natural experiment² on Twitter to investigate two research questions: *How does identity anonymity affect the quantity and quality of crowdsourced fact-checking on a social media platform? How do the characteristics of the crowdsourced users moderate the main impact?* Answering our research questions sheds light on how identity anonymity affects the crowdsourced fact-checking contribution and its heterogeneous effects. Theoretically, we enrich the understanding of the role of identity anonymity in the context of crowdsourced fact-checking. Furthermore, we provide empirical evidence of the impact of identity anonymity. Practically, this research provides managerial insights for social media platforms regarding best practices in identity anonymization policies; we hope to better regulate the misinformation spread.

In the following sections, we introduce related literature, followed by hypothesis development and research model. Then, we describe the research context and data, followed by an estimation strategy. Finally, we discuss empirical results, followed by a discussion and conclusion.

¹ <https://twitter.com/i/flow/join-birdwatch>

² In Nov. 22 of 2021, Twitter birdwatch program rolled out their user (fact-checkers) identity anonymization policy by making users use randomly generated alias, anonymizing users' fact-checking identity by deviating it from their social media identity.

Related Literature

Misinformation and Fact-checking

Misinformation comprises incorrect information, false claims, misleading statements, and false rumors (Pennycook et al. 2020). To combat misinformation, the industry has used several fact-checking methods, including third-party human fact-checkers (e.g., Snopes.com³), AI-based fact-checkers (e.g., Logically.ai⁴), crowdsourced fact-checkers (e.g., Twitter birdwatch program (Coleman 2021)), and human-AI combined approach (e.g., Poligraph proposed by Shan et al. (2021)) to flag misinformation, and promote nudges (e.g., Twitter nudging users to credible information⁵).

Several scholars have examined the efficacy of fact-checking methods. A group of studies (Figl et al. 2019; Moravec et al. 2020; Moravec et al. 2019; Ng et al. 2021; Pennycook et al. 2020; Ross et al. 2018) used behavioral experiments or field data to investigate the effectiveness of adopting flags from users (e.g., Pinocchio rating) or third-party human experts. They found that flagging a post can effectively mitigate user engagement and slow the spread of misinformation on social media. Several methods have been used to examine the effectiveness of crowdsourced flagging of misinformation (Allen et al. 2021a; Hassan et al. 2019; Pennycook and Rand 2019; Pinto et al. 2019). Using various types of methods (both qualitative and quantitative) on both experimental and field data, they found that crowdsourced users can be effective in fact-checking and flagging misinformation on social media, and crowdsourced fact-checking can be scaled up with a fast speed. Finally, studies have also investigated the effectiveness of using nudges and social norms (Gimpel et al. 2021; Hwang and Lee 2021) and found that nudging users toward credible information on Twitter can reduce the spread of misinformation and using social norms (both injunctive and descriptive social norm messages) can lead to the most misinformation reporting on social media platforms.

Though the literature has documented the effectiveness of fact-checking, especially crowdsourced fact-checking, little is known about how providing anonymity to fact-checkers impacts the quantity (e.g., contribution frequency) and quality (e.g., contribution trustworthiness and helpfulness) of fact-checking. In this research, we leverage the Twitter Birdwatch program policy change to examine how identity anonymity among the crowdsourced fact-checking affects users' contribution quantity and quality. Also, we investigate how the characteristics of the users moderate the impact of identity anonymity on their fact-checking contribution quantity and quality.

Crowdsourcing, Identity Anonymity, Online Disinhibition, and Social Presence

Crowdsourcing has been used for many applications, including healthcare (Khwanngern et al. 2019), journalism (Aitamurto 2016), task and solution-seeking (Mo et al. 2021), and user-generated content (Chen et al. 2017). For example, Chen et al. (2017) have investigated the motivating factors in voluntary contribution to a crowdsourcing knowledge contribution community via constructing a hidden Markov model. Crowdsourced fact-checking is when community users generate the fact-checking content for a news story within the community.

During the crowdsourced user-generated content review process, identity can affect the crowdsourced users' contribution quantity and quality. Studies have found that identity anonymity can make the users feel less concerned about being retaliated in reporting disruptive events such as flagging misinformation on social media (Lowry et al. 2013; Wong et al. 2021). In this situation, users feel less restraint in contributing to fact-checks, making the online disinhibition effect bigger. Thus, it is possible that identity anonymity can make crowdsourced users fact-check and flag misinformation more frequently.

However, anonymizing users' identities can also inhibit them from gaining more reputation, thus decreasing their contribution (Huang et al. 2017; Wang 2010). Furthermore, based on (Huang et al. 2017), identity disclosure can increase the volume of the user-generated content because disclosing identity can increase the social presence, improving social benefits and reputational gains via contributing more

³ <https://www.snopes.com/>

⁴ <https://www.logically.ai/>

⁵ <https://help.twitter.com/en/rules-and-policies/medical-misinformation-policy>

content. However, Pu et al. (2020) found that disclosing identity can inhibit users' willingness to generate content, leading to less volume of content generation.

Users' identity disclosure or anonymity can also affect the quality of the content they contribute. Based on (Pu et al. 2020), disclosing users' identities can increase their social presence, making them spend more time improving the generated content quality. Also, based on (Huang et al. 2017), increased social presence can make the users tend to express more emotions and use less negative tones in their posted content because the users' emotions are activated during a social environment, and they tend to build a positive social identity under identity disclosed.

Though literature underscores identity's impact on the crowdsourced users' contribution, the direction of the impact of identity anonymization in crowdsourced fact-checking is an open empirical question. Also, it is unknown whether and how the users' identity affects crowdsourced fact-checking. Fact-checking is the process of identifying whether a claim or news story is factual (Walter et al. 2020). It is time-consuming and needs expertise and understanding of the background of various topics. Fact-checkers also need to write the reports (e.g., notes in the Twitter Birdwatch) to justify the degree of certainty. Furthermore, news is a special goods that is the information about the events or activities, and news consumption is a socially-driven activity (Purcell et al. 2010). Therefore, crowdsourced fact-checking is different than other user-generated content (e.g., static writing online reviews), which makes this research important from a theoretical perspective.

Crowdsourced Fact-checking Contribution Performance

Crowdsourced fact-checking has been implemented by the industry (e.g., Twitter Birdwatch program) and investigated in academic research (Allen et al. 2021a; Godel et al. 2021; Micallef et al. 2020; Pennycook and Rand 2019; Pinto et al. 2019). Compared with third-party fact-checkers, crowdsourced fact-checkers can be scaled up and fast to mitigate the spread of misinformation on social media. However, their performance (e.g., fact-checking quality) depends on crowdsourced users' experience and domain knowledge, making it hard to perform at the same level as professional fact-checkers (Allen et al. 2021a; Godel et al. 2021).

Crowdsourced fact-checking contribution performance can be measured along two dimensions - quantity and quality. Quantity is an indicator of how crowdsourced fact-checking contributions can be scaled-up. Most of the misinformation on social media can be fact-checked quickly by their community users via crowdsourced fact-checking. Quality, which can be measured as helpfulness, and content trustworthiness have been documented as performance measurements in the crowdsourced fact-checking contribution (Allen et al. 2021b). In addition, belief and trust in the fact-checking flags can affect users' believability and engagement with the news (Moravec et al. 2019). Thus, both the helpfulness and trustworthiness of fact-checking should be considered a quality measure.

Hypothesis Development and Research Model

In this section, we propose our research model (shown in Figure 1) and hypothesize the effects of users' identity anonymity on crowdsourced fact-checking quantity and quality. First, we hypothesize the impact of identity anonymity on the fact-checking contribution quantity and quality using the theory of the social presence and the disinhibition effects. Then, we hypothesize the heterogeneous effect of the identity anonymity from the fact-checkers tenure perspective because crowdsourced users with different tenures have distinct experiences in fact-checking, making both the social presence and disinhibition effects have different impacts.

Social Presence and Disinhibition Effect

Within the crowdsourced fact-checking community, anonymizing the fact-checkers identity could reduce the social presence of the users. Literature has documented that user-generated content (e.g., online reviews) under social presence can lead to potential social benefits or reputational gains (Huang et al. 2017; Zhang and Zhu 2011). Thus, the loss of the social presence can decrease users' fact-checking contribution because they tend to gain fewer social benefits under identity anonymity. However, the countervailing mechanism may also exist. The loss of the social presence can make the users feel less concerned about being retaliated in reporting disruptive events (e.g., misinformation) on social media (Lowry et al. 2013; Wong et al. 2021).

Also, based on (Scott and Orlikowski 2014), identity anonymity can induce the online disinhibition effect, making fact-checkers feel more secure and comfortable contributing. Thus, identity anonymization can also increase the crowdsourced users' fact-checking contribution frequency. Accordingly, we propose two competing research hypotheses as follows.

H1a: *Identity anonymity reduces the crowdsourced users' fact-checking contribution frequency.*

H1b: *Identity anonymity increases the crowdsourced users' fact-checking contribution frequency.*

The identity anonymity can also affect the quality of the fact-checking contributions. On the one hand, the loss of social presence makes fact-checker gain less from the social benefits or reputational gains. On the other hand, the disinhibition effect makes them tend to use the language they would not use if their identity were not anonymized. Therefore, because of the loss of the social presence, crowdsourced users may tend to produce less helpful and trustworthy fact-checking content. However, users don't need to be concerned about others' evaluations due to the disinhibition effect. On the contrary, they tend to freely provide evidence and arguments to support their opinion, making fact-checking more helpful and trustworthy (Pu et al. 2020). Those competing mechanisms lead us to propose the following two competing hypotheses.

H2a: *Identity anonymity makes fact-checking more helpful and more trustworthy.*

H2b: *Identity anonymity makes fact-checking content less helpful and less trustworthy.*

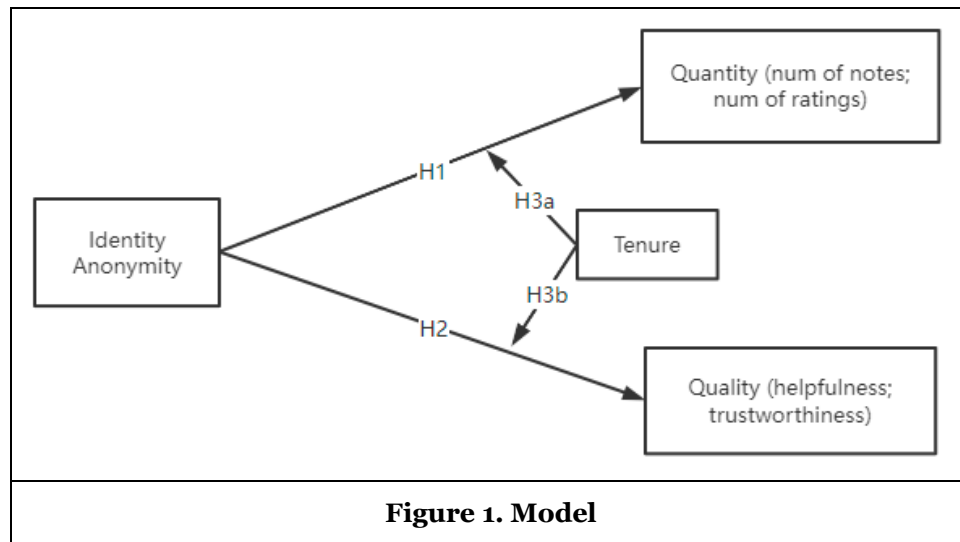
Heterogenous Effect - Tenure

The impact of identity anonymity on the fact-checking contribution frequency and quality can be heterogeneous for different types of users, especially when the users have different contribution experiences (e.g., with different tenure). The literature has documented the importance of individual characteristics in affecting the user-generated contribution quantity and quality. For example, Chen et al. (2017) revealed that individual characteristics such as tenure could affect users' motivation to contribute the content generation in the online Q&A community. In addition, Dong et al. (2020) adapted the status theory to the online community. They found that users' characteristics (community status standing) moderated the impact of users' self-interest and prosocial motivation on the post. Also, they found that users' tenure can affect users' contributions. Furthermore, literature documented that individual characteristics, including the community tenure, affected the user's contribution to the innovation community (Yan et al. 2018). For example, tenure indicates different experiences of users in generating content in the online community (Chen et al., 2017; Dong et al., 2020).

Since users' tenure, indicating users' different experience in making contribution, we anticipated the identity anonymity affects users' contribution behavior differently for users with different experience for several reasons. First, the longer tenure means the users have existed on the platform for a longer time, making them gain enough social presence from contributing. Thus, the tenure can moderate the identity anonymity impact of moderating the role of the gained social presence. Second, the more experiences of contributing brought by higher level of tenure can also make users form a regular habit in contributing, making them generate automatic contributing behavior (Kim et al. 2005) and less sensitive to policy changes such as identity anonymity. Third, the more experience of contributing (e.g., longer tenure) may make users feel more committed to contribute (Bateman et al. 2011), which can weaken the impact of identity anonymity. Users having different experiences in fact-checking news may have different attitudes (e.g., moderating the role of perceived social presence, forming habit and commitment in contributing) toward identity anonymization. The experiences may magnify or weaken the social presence and online disinhibition effects, thus moderating the impact of identity anonymization on the fact-checking contribution performance. Thus, we made the following research hypotheses.

H3a: *The impact of identity anonymity on the fact-checking quantity can be moderated by the contributor's tenure in the community.*

H3b: *The impact of identity anonymity on the fact-checking quality can be moderated by the contributor's tenure in the community.*



Research Context and Data


Research Context

We utilize Twitter Birdwatch (<https://twitter.com/i/birdwatch>) as the context to investigate our research questions and research model. Twitter launched the Birdwatch program on January 25, 2021, aiming to address misinformation by using crowdsourced users on the platform. They advocate Twitter users fact-check and flag the misinformation they meet on Twitter. Also, Twitter users are encouraged to write notes, providing informative context and explanations regarding the misinformation of a news article on Twitter. Furthermore, the program participants can also evaluate the quality of the fact-check notes from others by giving ratings such as helpfulness and trustworthiness.

One example of fact-checking in the news article "MyPillow CEO Mike Lindell rejoins Twitter just to be re-suspended" is put in Figure 2. Basically, the Birdwatch crowdsourced user with an anonymous note author name "Efficient Park Pigeon" provided the fact-checking notes in Figure 2(b) for the article in Figure 2(a). His fact-checking note reads as "Elon Musk", which is rated as helpful by other Birdwatch users. This fact-checking happened on May 1, 2022. Also, other users on the Twitter Birdwatch platform can rate this note as trustworthy and not trustworthy, which is stored in the database.

On November 22, 2021, Twitter introduced mandatory aliases where the Twitter username was replaced by an alias name (like the note author "Efficient Park Pigeon" in Figure 2(b)) in the Birdwatch program. This allowed users to flag and write notes for misinformation without revealing their real Twitter username.⁶ Users have to select an auto-generated alias (from among a few choices), which will protect and anonymize the users' identity, depriving from their public Twitter accounts. Also, the contribution name from the existing users before the policy change was replaced by their alias name. This exogenous identity anonymization event where all fact-checkers were made to adopt an alias instead of their Twitter username provides a natural experiment on how anonymity affects users' crowdsourced contribution quantity and quality.

⁶ This new function will make both the existing and new users use the auto-generated alias.

 <p>Jack Posobiec @JackPosobiec · 18h MyPillow CEO Mike Lindell rejoins Twitter one year after he was banned and days after Elon Musk bought the social media giant - but his account is suspended within four hours</p> <p>daily mail.co.uk MyPillow CEO Mike Lindell rejoins Twitter just to be re-suspended MyPillow CEO Mike Lindell briefly rejoined Twitter Sunday afternoon after being banned from the site in January 2021 for promoting claims ...</p>	<p>★ Currently rated helpful · Shown on Twitter · 18h Provides important context · Cites high-quality sources</p> <p>Elon Musk will not officially acquire or run Twitter, Inc. until the deal closes at what is projected to be October 24, 2022. No transaction has happened yet.</p> <p>https://www.republicworld.com/technology-news/social-media-news/elon-musks-twitter-buyout-deal-could-be-terminated-if-not-closed-by-october-24-articleshow.html#:~:text=Elon%20Musk%20will%20have%20to,6%20months%20may%20be%20provided.</p> <p>The deal is still in the “agreement” phase as operations are still being ran by the current regime.</p> <p>https://about.twitter.com/en/who-we-are/our-company</p> <p>× Note Details</p> <p>Current Status</p> <p>★ Helpful · Shown on Twitter Enough contributors from different perspectives agreed that this note is helpful, so it's being shown as context on the Tweet. Learn more</p> <p>Top tags selected by raters</p> <ul style="list-style-type: none"> Provides important context Cites high-quality sources <p>Note Author</p> <p>Efficient Park Pigeon</p>
(a) Twitter News Article	(b) Birdwatch Fact-checking Notes
Figure 2. An Example of Fact-checking Notes on Twitter Birdwatch	

Data

In this research, we focus on the user contribution data one month before and one month after the identity anonymization policy launched by the Twitter Birdwatch program. Twitter Birdwatch program is a community-based fact-checking of flagging misinformation on Twitter.⁷ Our data were drawn directly by the Twitter Birdwatch program from October 22, 2021, and Dec.22, 2021.⁸ Among our data, we found there are 1,291 users involved in fact-checking during this time. Also, we focus on the number of fact-checking notes and ratings as the contribution quantity measures and the note helpfulness ratio and trustworthiness ratio as the quality measures. Therefore, our data is composed of those variables shown in Table 1. To understand our data, we also conduct data summary statistics (shown in Table 2) and the variable correlation analysis (shown in Table 3).

Variables	Definitions
$Identity_i$	It represents whether the identity of user i is anonymized. = 1, Yes; = 0, No. This is the exogenous shock in our model.
$Tenure_i$	It represents the contribution tenure of user i on Birdwatch, measured by the number of days passed since the first contribution on Birdwatch until Nov. 22, 2021.
$Note_{it}$	It represents the number of fact-checking notes made by user i on a specific day t . This is a measure of the number of Twitter posts fact-checked by user i on day t .

⁷ https://blog.twitter.com/en_us/topics/product/2021/introducing-birdwatch-a-community-based-approach-to-misinformation

⁸ We only focused one month in the main analysis. In our future study, we will also try to use two months before and after the policy.

<i>Rating_{it}</i>	It represents the number of fact-checking evaluations made by user <i>i</i> on fact-checking notes made by other users on a specific day <i>t</i> . This is a measure of how much user <i>i</i> evaluates the fact-checking effort of other users.
<i>Helpfulness_{it}</i>	It represents the ratio of the number of fact-checking notes made by user <i>i</i> on day <i>t</i> , which are rated as helpful by others (as a fraction of the total number of fact-checking notes made by user <i>i</i> on day <i>t</i>).
<i>Trustworthy_{it}</i>	It represents the ratio of the number of fact-checking notes made by user <i>i</i> on day <i>t</i> , which are rated as trustworthy by others (as a fraction of the total number of fact-checking notes made by user <i>i</i> on day <i>t</i>).
Table 1. Variables and Definitions	

Note: the unit of analysis of this study is a user (fact-checker) and day. In other words, all our variables are measured at the daily level for each user (fact-checker) in our sample. We use the ratio to measure helpfulness and trustworthiness because the count of the helpful and trustworthy notes can be biased toward the total number of notes per day.

Variables	Obs	Mean	Std.Dev.	Min	Max
Identity (1)	2,741	.44	.50	0	1
Tenure (2)	2,741	22.54	9.22	1	31
Notes (3)	2,741	1.48	1.92	1	36
Ratings (4)	2,741	1.82	4.93	0	92
Helpfulness (5)	2,741	.74	.33	0	1
Trustworthiness (6)	2,741	.74	.43	0	1
Table 2. Variable Summary Statistics					

Variables	(1)	(2)	(3)	(4)	(5)	(6)
Identity (1)	1.00					
Tenure (2)	-0.21	1.00				
Notes (3)	-0.02	0.14	1.00			
Ratings (4)	-0.02	0.13	0.21	1.00		
Helpfulness (5)	-0.03	-0.03	-0.03	-0.03	1.00	
Trustworthiness (6)	0.06	0.12	0.10	0.07	0.17	1.00
Table 3. Variable Correlation Matrixes						

Note. The values in Table 3 represent the Pearson correlation coefficient.

From Table 2, the mean value of anonymity is 0.44, indicating that 44% of observations happen after the identity is anonymized. The mean value of the tenure is 22.54, indicating the mean values of the users exist for 22.54 days, on average, before the identity anonymization policy. Table 2 also shows that the average number of fact-checking notes and their evaluations (ratings) is 1.48 and 1.82 daily. The average note helpfulness and trustworthiness ratios are both 0.74, indicating that 74% of daily notes are rated helpful, and 74% are rated as trustworthy, on average, respectively. From Table 3, we can know all of the variables considered have small correlations (the absolute coefficient is less than 0.3), following the guidelines in the literature (Cohen 1988).

Estimation Strategy

We use the regression discontinuity in time (RDiT) design to estimate the average treatment effects of identity anonymity on the crowdsourced fact-checking contribution frequency and quality. RDiT is a type of regression discontinuity (RD) design where time is the running variable (Hausman and Rapson 2018). RD is a widely used research design to estimate the causal effects of inventions where the assignment of the treatment is determined by an observed running variable (Imbens and Lemieux 2008). The RDiT

identification strategy fits well with our scenario and data because (1) the identity anonymity policy on the Twitter Birdwatch platform provides a natural/quasi-experiment invention, anonymizing all of the contributors' identities just on November 22, 2021; (2) the policy change date is clear and known; (3) the policy is discrete with perfect compliance.

RD utilizes a quasi-experimental approach to estimate the local treatment effects. The fundamental principle of using RD is to find the cut-off value of an observed running variable and then organize the control group and treatment groups using the just below and just above neighborhoods of the cut-off point (Cattaneo et al. 2019; Imbens and Lemieux 2008; Lee and Lemieux 2010). The two neighborhoods around the cut-off point are so similar that the two groups are as good as randomly assigned. The observed running variable, also called as assignment variable, can be many things such as geo-location, testing scores, and time of an event (Kolesár and Rothe 2018). We use the identity anonymity policy time as the running variable in this research.

Furthermore, RD can be sharp and fuzzy, depending on the treatment assignment (Imbens and Lemieux 2008). The RD is sharp if the treatment is a deterministic and discontinuous function of the running variable. Otherwise, if the treatment is probabilistic of the function of the running variable, the RD is fuzzy. In this study, we use the sharp RD because the treatment status (anonymized) is deterministic and discontinuous regarding the running variable, time (after the policy, the treatment is assigned; before the policy, the treatment is not assigned). To estimate the anonymity treatment effects under RD, we capture the outcome trend before and after the anonymity policy. We use the polynomial regression function with covariates to estimate the treatment effects (Calonico et al. 2017; Calonico et al. 2014). This approach is general and flexible that it avoids needing to precisely capture the regression trend on either side of the discontinuity and the biased kernel estimators at boundaries (Imbens and Kalyanaraman 2012; McCrary 2008). However, it needs to be careful in selecting a proper bandwidth, function form (e.g., different number of the polynomial orders), and kernel function (Cattaneo et al. 2019). Thus, we need to conduct a group of robustness checks around the selections.

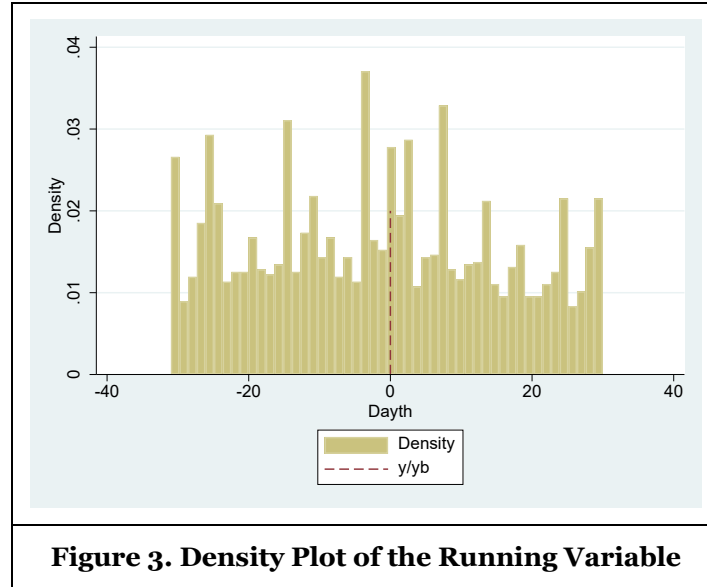
Aside from the model and parameters, there are also some issues around the selected time running variable. Using time as the running variable can bring the sorting problem. The subjects can purposely behave differently around the discontinuity point, which asks for the density test around the point (Hausman and Rapson 2018; McCrary 2008). The Twitter Birdwatch announced the identity anonymization policy, which was inherently exogenous. The current and new users on the platform will automatically anonymize their identity by using an automatically generated alias. Therefore, the sorting problem should not exist as the users could not predict the identity anonymization policy and change their behavior around the policy announcement time. To confirm this argument, we do a group of manipulation tests.

Furthermore, using time as the running variable, we can also see the seasonality and autoregressive issues of the regression. Thus, we conduct robustness checks to deal with them. We follow the whole process of testing the validity of our RD (Imbens and Lemieux 2008). Furthermore, we are also interested in exploring the heterogeneous effect of identity anonymization. The standard RD couldn't support examining the moderators by including an interaction term. Thus, to test the moderation effect of the tenure, we add an interaction term of the tenure and anonymity policy while conducting the regression discontinuity (RD) analysis, following Cortes et al. (2015) and Pustejovsky (2016)

Empirical Results

Manipulation Test

We first ensure the running variable (time in our study) is not manipulated that the fact-checking contributors enrolling in the Twitter Birdwatch program could not self-select to receive the treatment (identity anonymization in this research). Then, to further understand whether our collected data doesn't contain the running variable manipulation issue, we first checked whether the density of the running variable is continuous around the cut-off point by plotting the histogram of the running variable. The result is shown in Figure 3, where we didn't see a clear discontinuity around the cut-off point. Furthermore, we conduct a formal manipulation test using the *rddensity* STATA command (with its default setting), following Cattaneo et al. (2018). The results are shown in Table 4.



C = 0.000	Left of c	Right of c
Number of obs	1433	1308
Eff. Number of obs	1044	1050
Order est. (p)	2	2
Order bias (q)	3	3
BW est. (h)	23.000	23.000

Table 4. RD Manipulation Test using Local Polynomial Density Estimation

Note. Number of obs = 2741; Model = unrestricted; BW method = comb; Kernel = triangular; VCE method = jackknife; Running variable is time; $T_q(h_{l,comb,p}, h_{r,comb,p}) = -0.6941$ and p -value = 0.4876.

The regression discontinuity manipulation test results reveal that it doesn't have sufficient evidence to argue the running variable is manipulated with the test statistic value of -0.6941 and p -value of 0.4876 (>0). Therefore, our data pass the manipulation test. Also, Table 4 reveals 1433 and 1308 observations in the control and treatment groups. Using the bandwidth as (23.000, 23.000) for an unrestricted model with $q = 3$ polynomial and polynomial order $p = 2$, we can get 1044 and 1050 effective samples in the control and treatment groups respectively.

Estimation Results of the Main Effect

To test our research model, we first investigate how identity anonymization affects fact-checking contribution quantity and quality. Following Cattaneo et al. (2019), we estimate the local causal effect of the treatment on the outcome variables. To visually understand how identity anonymization affects our different outcome variables, we plot a 4th degree polynomial fit to the entire sample. The results are shown in Figure 4. It indicates that identity anonymization may decrease the contributed number of notes but increase the ratio of the helpful and trustworthy notes. For the number of ratings, identity anonymization may not affect it. We conduct the RD analysis using rd STATA command to quantify the causal impact (Nichols 2011). To test the robustness of the results, we also set the bandwidths as one/two/three weeks (7/14/21 days). The results are shown in Table 5.

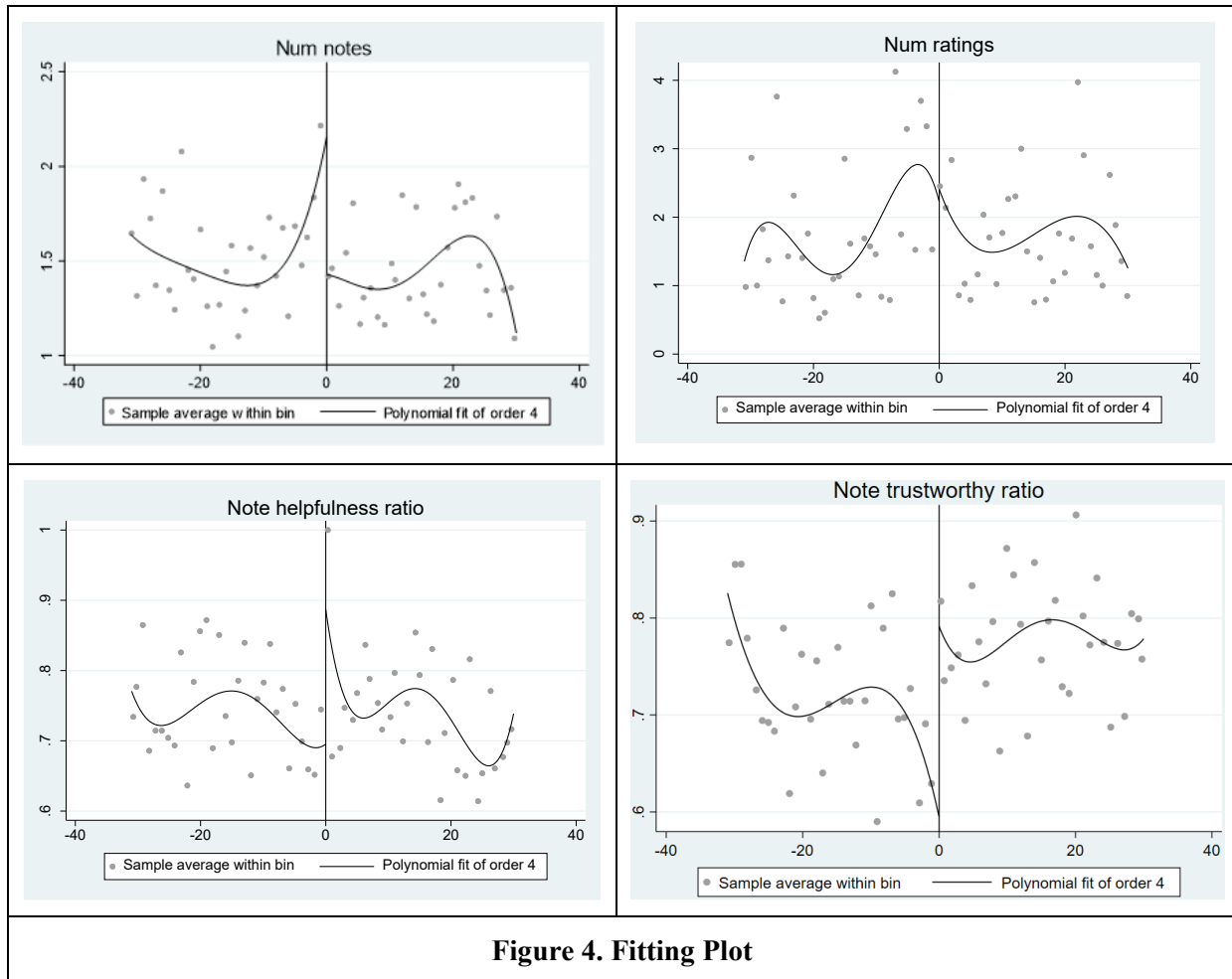


Figure 4. Fitting Plot

Bandwidth	Notes	Ratings	Helpful note ratio	Trustworthy note ratio
7	-0.867 (.51)	.990 (.78)	.179*** (.05)	.165* (0.07)
14	-.519 (.29)	-.236 (.65)	.169*** (.03)	.138** (0.05)
21	-.402* (.19)	-.363 (.58)	.142*** (.03)	.115** (0.04)

Table 5. RD Estimation Results of Identity Anonymization

Note. * $p < .05$, ** $p < .01$, *** $p < .001$; Standard error is in the brackets; Estimators are in the box with standard errors in the parenthesis. We used the default triangle kernel; We have tried the epanechnikov and uniform kernel functions (Calonic et al. 2017) and found robust results.

From Table 5, identity anonymization significantly reduces the number of notes (only for using the bandwidth at 21 days), but increases the helpful note ratio and trustworthy note ratio. Using the bandwidth as seven days, the local treatment effects of the identity anonymization on the helpful and trustworthy note ratios are 0.179 and 0.165, indicating that anonymizing the fact-checker's identity by making them using an alias on the Birdwatch platform can increase the helpful note ratio by 0.179, and increase trustworthy note ratio by 0.165 daily, on average.

Regarding RQ1 and H1, and H2, our empirical results confirm that identity anonymity can affect crowdsourced fact-checking contribution. However, our results only support the impact of identity

anonymization on the fact-checking quality (H2 is supported) rather than quantity (H1 is not supported). Though the impact on the number of notes can be significant, it is sensitive to the bandwidth selection, making the results not robust. And, it can increase the fact-checking quality by increasing the helpful and trustworthy note ratios (H2b is supported).

There are two possible explanations for the impact of identity anonymity on the fact-checking quantity. It is possible that fact-checking users don't perceive the social presence and online disinhibition matter as affecting their contribution frequency. They contribute the fact-checking ratings and notes for other reasons, such as informing and spreading facts on social media rather than increasing the social presence or avoiding online inhibition. Thus, the anonymizing identity doesn't affect the contribution quantity. Besides, it may also be true that both the social presence and online disinhibition effects affect users' fact-checking contribution quantity. But, it is likely that anonymizing the users' identity makes both effects countervail and make the overall impact not significant. Also, there are possible other alternative explanations of the identity anonymity on the fact-checking contribution quantity such as self-efficacy and perceived outcome effectiveness. However, the total effect of those mechanisms may mitigate with each other and then form a general non-impact on the fact-checking quantity.

For the impact of identity anonymity on the fact-checking quality, the disinhibition effect should be more salient than the social presence effect, or only the disinhibition effect exists, making users contribute more helpful and trustworthy fact-checking content. That indicates that crowdsourced users may feel less restricted in putting more controversial evidence and useful context to support their fact-checking.

Our results can benefit the managers of crowdsourced platforms. On the one hand, they can leverage the identity anonymization policy to increase users' contribution quality. But, on the other hand, they can implement solutions to increase the social presence (e.g., introducing communication among contributors such as allowing commenting on the fact-checking notes to increase the social presence effect (Pu et al. 2020)) to increase the contribution quantity.

To evaluate the robustness of our results, we also set different polynomial orders (from one to three) and used different kernel functions (triangular, epanechnikov, and uniform). The results shown in Table 6 reveal that the impact of identity anonymization on the number of notes is not robust, but the helpful note and trustworthy ratios are robust across different settings. Therefore, our research findings are further confirmed.

Setting	Notes	Helpful note ratio	Trustworthy note ratio
p (1), triangular	-0.465* (0.23)	0.142*** (0.03)	0.115** (0.04)
p (1), epanechnikov	-0.448* (0.21)	0.133*** (0.03)	0.108** (0.04)
p (1), uniform	-0.427* (0.19)	0.130*** (0.03)	0.0940* (0.04)
p (2), triangular	-0.623 (0.38)	0.187*** (0.04)	0.164** (0.06)
p (2), epanechnikov	-0.591 (0.36)	0.181*** (0.04)	0.158** (0.06)
p (2), uniform	-0.544 (0.31)	0.155*** (0.04)	0.144** (0.05)
p (3), triangular	-0.894 (0.59)	0.209*** (0.06)	0.203* (0.09)
p (3), epanechnikov	-0.841 (0.56)	0.213*** (0.06)	0.204* (0.08)
p (3), uniform	-0.719 (0.50)	0.219*** (0.05)	0.191* (0.08)

Table 6. Robustness Checks using different Polynomial Order and Kernel Function

Note. * $p < .05$, ** $p < .01$, *** $p < .001$; Standard error is in the brackets; the bandwidth used for this robustness is 21 days; we test the polynomial order from 1 to 3 and use three different kernel functions: triangular, epanechnikov, and uniform. Other settings are in default using the STATA `rdrobust` command (Calonico et al. 2017).

To test whether the impact of the identity anonymization on the fact-checking contribution is random rather than identity policy, we conduct a few falsification tests. Specifically, instead of treating the cut-off point, happening policy time as November 22, 2021, we selected the cut-off point as the before 20 and 10 and after 10 and 20 days. Then, we conduct the same regression discontinuity analysis. The results are shown in Table 7, revealing that identity anonymization doesn't affect the fact-checking contributions. Thus, the effect comes from identity anonymization, which is not a random effect.

Cut-off points	Notes	Helpful note ratio	Trustworthy note ratio
-20	-0.170 (0.25)	0.0755 (0.04)	0.0780 (0.05)
-10	0.295 (0.18)	-0.00213 (0.03)	0.00853 (0.04)
10	0.284 (0.17)	-0.00432 (0.03)	0.0493 (0.04)
20	0.493 (0.28)	-0.0154 (0.04)	0.0571 (0.05)

Table 7. Falsification Tests on Different Cut-off Points

Note. * $p < .05$, ** $p < .01$, *** $p < .001$; Standard error is in the brackets; the bandwidth used for this robustness is 21 days; we test the cut-off points among -20/-10/10/20 with other settings as default using the STATA `rdrobust` command (Calonico et al. 2017).

Estimation Results of the Tenure Effect

To investigate how the tenure of the fact-checking users' moderates the main impact, we focus on both the fact-checking contribution quantity and quality outcomes. Though the main impacts on the quantity are not significant, conducting the moderation analysis could provide mechanisms to better understand the identity anonymization impact. By adding the interaction term of the identity anonymization and the tenure into the RD analysis, we show the results in Table 8.

	Notes	Ratings	Helpful note ratio	Trustworthy note ratio
Running variable	0.012*	0.023	-0.003***	0.000
	(0.01)	(0.01)	(0.00)	(0.00)
Identity	0.158	-0.158	0.018	0.174**
	(0.18)	(0.44)	(0.04)	(0.01)
Tenure	0.051***	0.070***	-0.004***	0.001***
	(0.01)	(0.02)	(0.00)	(0.00)
Identity*Tenure	-0.020*	-0.001	0.004*	-0.000
	(0.01)	(0.02)	(.00)	(0.00)
Constant	0.459***	0.475	0.795***	0.487***
	(0.12)	(0.32)	(0.03)	(0.05)

Table 8. Moderation Impact of the Tenure

Note. * $p < .05$, ** $p < .01$, *** $p < .001$; Standard error is in the brackets.

The interaction of identity and tenure significantly affects the number of the notes, and the helpful note ratio. It indicates that the users' tenure can moderate the impact of identity anonymization on the number of notes, and helpful note ratio. Specifically, the longer the tenure is (the users exist in this platform longer), the bigger the impact of identity anonymization on the helpful note ratio. Based on the results, if the users exist on the platform for an additional day, the effect of identity anonymization on the helpful note ratio could be 0.004 higher.

Regarding the RQ2 and research H3, we find that tenure can moderate the impact of identity anonymization on the fact-checking quantity and quality. However, we need to conduct further analysis to understand how identity anonymization affects the quantity, specifically for the number of the notes, and we only find that the tenure affects the impact on the helpful note ratio rather than other measures. This finding can benefit the fact-checking platforms which use crowdsourced fact-checking such as Twitter Birdwatch. Their identity anonymization policy will benefit more in improving the trustworthy note ratios if they can release it after they accumulate users with longer tenure.

Discussion and Conclusion

According to Harvard Business Review, the spread of misinformation on social media platforms (Acemoglu et al. 2021) has been called a crisis and needs to be fixed (Azhar 2021). Both the industry and the academy have been trying various methods to slow the spread of misinformation. Among them, community-based

crowdsourced fact-checking has been confirmed as an effective and fast method to flag misinformation and slow its spread.

Our study is the first to investigate how identity anonymization affects the crowdsourced users' contribution, especially in the fact-checking context. We found that anonymizing users' identity will not affect their fact-checking frequency regarding the number of notes and the number of ratings. However, it does improve the contribution quality by improving the helpful and trustworthy note ratios. Those findings confirm the existence of the online disinhibition effect. Also, it is possible that the social presence also affects the user's fact-checking contribution. Because of their countervailing mechanisms, identity anonymization doesn't affect the fact-checking contribution frequency. However, the online disinhibition effect should be more salient in the contribution quality because the helpful and trustworthy fact-checking note ratios increase after the identity anonymity is activated. We summarize the results and findings of the research hypothesis testing in Table 9. Those findings reveal that crowdsourced fact-checking users do care about online disinhibition. They may tend to avoid checking controversial news when their identity is disclosed because of feeling they are retaliated, which has been confirmed by Poynter reports.⁹

Research Hypotheses	Findings	Results
H1a: <i>Identity anonymity reduces the crowdsourced users' fact-checking contribution frequency.</i>	Identity anonymity doesn't affect the contribution frequency	Neither competing hypothesis supported
H1b: <i>Identity anonymity increases the crowdsourced users' fact-checking contribution frequency.</i>	Identity anonymity doesn't affect the contribution frequency	
H2a: <i>Identity anonymity makes fact-checking more helpful and more trustworthy.</i>	Identity anonymity increases the helpful and trustworthy note ratios	H2b Supported
H2b: <i>Identity anonymity makes fact-checking content less helpful and less trustworthy.</i>	Identity anonymity increases the helpful and trustworthy note ratios	
H3a: <i>The impact of identity anonymity on the fact-checking quantity can be moderated by the contributor's tenure in the community.</i>	The tenure moderates the impact of the identity anonymity on the number of the notes	Partially supported
H3b: <i>The impact of identity anonymity on the fact-checking quality can be moderated by the contributor's tenure in the community.</i>	The tenure moderates the impact of the identity anonymity on the helpfulness ratio	Partially supported
Table 9. Results of the Hypothesis Testing		

Our research makes theoretical contributions and practical implications. First, our research enriches the IS literature by investigating the identity anonymity among the crowdsourcing community. We are among the first to identify how identity anonymization affects the crowdsourced users' contribution quantity and quality. Second, our research also enriches the IS literature around misinformation and fact-checking. We are among the first to understand the crowdsourced fact-checking dynamics from the angle of identity anonymization and its moderation impact of the users' tenure.

Our research also has a few practical implications. On the one hand, we bring managerial insights to the social media platforms. We show that the identity anonymization policy may not increase fact-checking users' contribution quantity, but the fact-checking quality does increase. Thus, we call for their attention to take some measures to increase the contribution frequency, such as increasing the social presence existence. On the other hand, we suggest that social media platforms aiming to adopt the community-based crowdsourced fact-checking consider the timing to adopt the identity anonymization policy. Based on our

⁹ <https://www.poynter.org/fact-checking/2018/these-fact-checkers-were-attacked-online-after-partnering-with-facebook/>

findings, with fact-checking users having higher tenure, the benefits of the identity anonymization on the trustworthy fact-checking notes will be more pronounced.

While this research explores the crowdsourced fact-checking dynamics around identity anonymization, many areas remain to be explored. Future studies might build on this work and focus on testing the underlying mechanisms of identity anonymization on the fact-checking contribution quantity and quality. Moreover, it deserves to explore more moderating factors from the angle of the fact-checkers' demographics such as gender, race, and education background or the aspect of the news categories. Those examinations can better establish the understanding of the dynamics of the community-based crowdsourced fact-checks.

Future Research

For our next step, we first tend to explore more moderating effects, focusing on factors such as the experience and quality of the crowdsourced fact-checkers. We will also conduct a sub-group analysis to understand further how identity anonymity affects the contribution quantity in different sub-groups. Furthermore, we may also design online lab experiments to directly measure and test the underlying mechanisms explaining the impact of identity anonymity. Also, to furtherly confirm our research findings are robust, we plan to conduct a few robustness checks. For example, we plan to enrich the selected data range from two months to four months. And we plan to conduct the Twitter trending topic analysis to remove the potential confounders of the impact of a sudden major issue during our selected time range. Besides, we plan to check the traffic difference before and after the policy change in order to rule out the possibility that the observed effects are caused by the general traffic difference. Furthermore, to firmly understand whether users' fact-checking quality will be changed for the controversial news after the identity anonymity, we plan to conduct some fact-checking quality comparisons before and after the policy change on the controversial news. Finally, we plan to add some additional covariates in our next step to furtherly increase the validity of our research design.

References

- Acemoglu, D., Ozdaglar, A., and Siderius, J. 2021. "Misinformation: Strategic Sharing, Homophily, and Endogenous Echo Chambers," National Bureau of Economic Research.
- Aitamurto, T. 2016. "Crowdsourcing as a Knowledge-Search Method in Digital Journalism: Ruptured Ideals and Blended Responsibility," *Digital journalism* (4:2), pp. 280-297.
- Allen, J., Arechar, A. A., Pennycook, G., and Rand, D. G. 2021a. "Scaling up Fact-Checking Using the Wisdom of Crowds," *Science advances* (7:36), p. eabf4393.
- Allen, J. N. L., Martel, C., and Rand, D. 2021b. "Birds of a Feather Don't Fact-Check Each Other: Partisanship and the Evaluation of News in Twitter's Birdwatch Crowdsourced Fact-Checking Program,").
- Azhar, A. 2021. "Fixing the social media crisis," *Harvard Business Review*, February 24 (available at <https://hbr.org/podcast/2021/02/fixing-the-social-media-crisis>; retrieved April 29, 2022).
- Bateman, P.J., Gray, P.H. and Butler, B.S., 2011. "Research note—the impact of community commitment on participation in online communities," *Information Systems Research* (22:4), pp.841-854.
- Burtch, G., Ghose, A., & Wattal, S. (2015). The hidden cost of accommodating crowdfunder privacy preferences: A randomized field experiment. *Management Science* (61:5), pp.949-962.
- Calonico, S., Cattaneo, M. D., Farrell, M. H., and Titiunik, R. 2017. "Rdrobust: Software for Regression-Discontinuity Designs," *The Stata Journal* (17:2), pp. 372-404.
- Calonico, S., Cattaneo, M. D., and Titiunik, R. 2014. "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs," *Econometrica* (82:6), pp. 2295-2326.
- Castillo, M. 2019. "Exclusive: Fake news is costing the world \$78 billion a year," *Cheddar*, Cheddar, November 18 (available at <https://cheddar.com/media/exclusive-fake-news-is-costing-the-world-billion-a-year>; retrieved April 29, 2022).
- Cattaneo, M. D., Idrobo, N., and Titiunik, R. 2019. *A Practical Introduction to Regression Discontinuity Designs: Foundations*. Cambridge University Press.
- Cattaneo, M. D., Jansson, M., and Ma, X. 2018. "Manipulation Testing Based on Density Discontinuity," *The Stata Journal* (18:1), pp. 234-261.
- Chen, W., Wei, X., and Zhu, K. 2017. "Engaging Voluntary Contributions in Online Communities: A Hidden Markov Model," *Mis Quarterly* (42:1), pp. 83-100.

- Cohen, J. 1988. "Set Correlation and Contingency Tables," *Applied Psychological Measurement* (12:4), pp. 425-434.
- Coleman, K. 2021. "Introducing Birdwatch, a community-based approach to misinformation," *Twitter*, Twitter, January 25 (available at https://blog.twitter.com/en_us/topics/product/2021/introducing-birdwatch-a-community-based-approach-to-misinformation; retrieved April 29, 2022).
- Cortes, K. E., Goodman, J. S., and Nomi, T. 2015. "Intensive Math Instruction and Educational Attainment Long-Run Impacts of Double-Dose Algebra," *Journal of Human Resources* (50:1), pp. 108-158.
- Dong, L., Huang, L., Hou, J. J., and Liu, Y. 2020. "Continuous Content Contribution in Virtual Community: The Role of Status-Standing on Motivational Mechanisms," *Decision Support Systems* (132), p. 113283.
- Figl, K., Kießling, S., Rank, C., and Vakulenko, S. 2019. "Fake News Flags, Cognitive Dissonance, and the Believability of Social Media Posts,").
- Gimpel, H., Heger, S., Olenberger, C., and Utz, L. 2021. "The Effectiveness of Social Norms in Fighting Fake News on Social Media," *Journal of Management Information Systems* (38:1), pp. 196-221.
- Godel, W., Sanderson, Z., Aslett, K., Nagler, J., Bonneau, R., Persily, N., and Tucker, J. A. 2021. "Moderating with the Mob: Evaluating the Efficacy of Real-Time Crowdsourced Fact-Checking," *Journal of Online Trust and Safety* (1:1).
- Hassan, N., Yousuf, M., Mahfuzul Haque, M., A. Suarez Rivas, J., and Khadimul Islam, M. 2019. "Examining the Roles of Automation, Crowds and Professionals Towards Sustainable Fact-Checking," *Companion Proceedings of The 2019 World Wide Web Conference*, pp. 1001-1006.
- Hausman, C., and Rapson, D. S. 2018. "Regression Discontinuity in Time: Considerations for Empirical Applications," *Annual Review of Resource Economics* (10), pp. 533-552.
- Huang, N., Hong, Y. K., and Burtch, G. 2017. "Social Network Integration and User Content Generation: Evidence from Natural Experiments," *MIS Quarterly* (41:4), pp. 1035-1058.
- Hwang, E. H., and Lee, S. 2021. "A Nudge to Credible Information as a Countermeasure to Misinformation: Evidence from Twitter," *Available at SSRN*.
- Imbens, G., and Kalyanaraman, K. 2012. "Optimal Bandwidth Choice for the Regression Discontinuity Estimator," *The Review of Economic Studies* (79:3), pp. 933-959.
- Imbens, G. W., and Lemieux, T. 2008. "Regression Discontinuity Designs: A Guide to Practice," *Journal of Econometrics* (142:2), pp. 615-635.
- Khwanngern, K., Natwichai, J., Sitthikham, S., Sitthikamtiub, W., Kaveeta, V., Rakchittapoke, A., and Martkamjan, S. 2019. "Crowdsourcing Platform for Healthcare: Cleft Lip and Cleft Palate Case Studies," *International Conference on Network-Based Information Systems*: Springer, pp. 465-474.
- Kim, S.S., Malhotra, N.K. and Narasimhan, S., 2005. "Research note—two competing perspectives on automatic use: A theoretical and empirical comparison," *Information Systems Research* (16:4), pp.418-432.
- Kolesár, M., and Rothe, C. 2018. "Inference in Regression Discontinuity Designs with a Discrete Running Variable," *American Economic Review* (108:8), pp. 2277-2304.
- Lee, D. S., and Lemieux, T. 2010. "Regression Discontinuity Designs in Economics," *Journal of Economic Literature* (48:2), pp. 281-355.
- Lowry, P. B., Moody, G. D., Galletta, D. F., and Vance, A. 2013. "The Drivers in the Use of Online Whistle-Blowing Reporting Systems," *Journal of Management Information Systems* (30:1), pp. 153-190.
- Lu, Y., Jerath, K., and Singh, P. V. 2013. "The Emergence of Opinion Leaders in a Networked Online Community: A Dyadic Model with Time Dynamics and a Heuristic for Fast Estimation," *Management Science* (59:8), pp. 1783-1799.
- Mahadevan, A., and Mantas, H. 2021. "Here are Twitter's most prolific citizen fact-checkers," *Poynter*, July 15 (available at <https://www.poynter.org/fact-checking/2021/here-are-twitters-most-prolific-citizen-fact-checkers/>; retrieved April 29, 2022).
- McCrary, J. 2008. "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test," *Journal of econometrics* (142:2), pp. 698-714.
- Micallef, N., He, B., Kumar, S., Ahamad, M., and Memon, N. 2020. "The Role of the Crowd in Countering Misinformation: A Case Study of the Covid-19 Infodemic," *2020 IEEE International Conference on Big Data (Big Data)*: IEEE, pp. 748-757.
- Mo, J., Sarkar, S., and Menon, S. 2021. "Competing Tasks and Task Quality: An Empirical Study of Crowdsourcing Contests," *MIS Quarterly* (45:4).

- Moravec, P. L., Kim, A., and Dennis, A. R. 2020. "Appealing to Sense and Sensibility: System 1 and System 2 Interventions for Fake News on Social Media," *Information Systems Research* (31:3), pp. 987-1006.
- Moravec, P. L., Minas, R. K., and Dennis, A. 2019. "Fake News on Social Media: People Believe What They Want to Believe When It Makes No Sense at All," *MIS Quarterly* (43:4), pp. 1343-1360.
- Ng, K. C., Tang, J., and Lee, D. 2021. "The Effect of Platform Intervention Policies on Fake News Dissemination and Survival: An Empirical Examination," *Journal of Management Information Systems* (38:4), pp. 898-930.
- Pennycook, G., Bear, A., Collins, E. T., and Rand, D. G. 2020. "The Implied Truth Effect: Attaching Warnings to a Subset of Fake News Headlines Increases Perceived Accuracy of Headlines without Warnings," *Management Science* (66:11), pp. 4944-4957.
- Pennycook, G., and Rand, D. G. 2019. "Fighting Misinformation on Social Media Using Crowdsourced Judgments of News Source Quality," *Proceedings of the National Academy of Sciences* (116:7), pp. 2521-2526.
- Pennycook, G., and Rand, D. G. 2021. "Examining False Beliefs About Voter Fraud in the Wake of the 2020 Presidential Election," *The Harvard Kennedy School Misinformation Review*.
- Pinto, M. R., de Lima, Y. O., Barbosa, C. E., and de Souza, J. M. 2019. "Towards Fact-Checking through Crowdsourcing," *2019 IEEE 23rd International Conference on Computer Supported Cooperative Work in Design (CSCWD): IEEE*, pp. 494-499.
- Pu, J., Chen, Y., Qiu, L., and Cheng, H. K. 2020. "Does Identity Disclosure Help or Hurt User Content Generation? Social Presence, Inhibition, and Displacement Effects," *Information Systems Research* (31:2), pp. 297-322.
- Purcell, K., Rainie, L., Mitchell, A., Rosenstiel, T., and Olmstead, K. 2010. "Understanding the Participatory News Consumer," *Pew Internet and American Life Project* (1), pp. 19-21.
- Pustejovsky, J. E. 2016. "Regression discontinuities with covariate interactions in the RDD package," James E. Pustejovsky, James E. Pustejovsky, January 25 (available at <https://www.jepusto.com/rdd-interactions/>; retrieved April 29, 2022).
- Rodgers, K., and Massac, N. 2020. "Misinformation: A Threat to the Public's Health and the Public Health System," *Journal of Public Health Management and Practice* (26:3), pp. 294-296.
- Ross, B., Jung, A., Heisel, J., and Stieglitz, S. 2018. "Fake News on Social Media: The (in) Effectiveness of Warning Messages," *39th International Conference on Information Systems: Association for Information Systems*, p. 16.
- Scott, S. V., and Orlikowski, W. J. 2014. "Entanglements in Practice," *MIS Quarterly* (38:3), pp. 873-894.
- Shan, G., Zhao, B., Clavin, J. R., Zhang, H., and Duan, S. 2021. "Poligraph: Intrusion-tolerant and distributed fake news detection system," *IEEE Transactions on Information Forensics and Security* (17), pp. 28-41.
- Von Krogh, G., Haefliger, S., Spaeth, S., and Wallin, M. W. 2012. "Carrots and Rainbows: Motivation and Social Practice in Open Source Software Development," *MIS Quarterly* (36:2), pp. 649-676.
- Walter, N., Cohen, J., Holbert, R. L., and Morag, Y. 2020. "Fact-Checking: A Meta-Analysis of What Works and for Whom," *Political Communication* (37:3), pp. 350-375.
- Wang, Z. 2010. "Anonymity, Social Image, and the Competition for Volunteers: A Case Study of the Online Market for Reviews," *The BE Journal of Economic Analysis & Policy* (10:1).
- Wong, R. Y. M., Cheung, C. M., Xiao, B., and Thatcher, J. B. 2021. "Standing up or Standing By: Understanding Bystanders' Proactive Reporting Responses to Social Media Harassment," *Information Systems Research* (32:2), pp. 561-581.
- Yan, J., Leidner, D. E., and Benbya, H. 2018. "Differential Innovativeness Outcomes of User and Employee Participation in an Online User Innovation Community," *Journal of Management Information Systems* (35:3), pp. 900-933.
- 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.