Interplay of network structure and neighbour performance in user innovation

Kunhao Yang1*, Itsuki Fujisaki1,2 & Kazuhiro Ueda1*

ABSTRACT Previous studies showed that regular users have become an important source of innovation (called user innovation). Previous studies also suggested that two factors have significant impacts on user innovation: the social network structure of the users and neighbours’ innovation performance (neighbours mean users having interactions with the focal user). However, in these studies, the influence of the two factors were only discussed separately, and it remained unclear whether these two factors interdependently affected a focal user’s innovation. To examine the interplay between the network structure and the neighbours’ innovation performance, we harnessed data sets from “Idea Storm”, which collects data on user network and idea submission. Through panel regression analyses, we found that—within an open-network structure—the higher innovation performance of neighbours has a larger positive impact on the focal user’s innovation ability. Conversely, in an enclosed network structure, neighbours’ higher performance has a larger negative impact on the focal user’s innovation ability. Our findings filled an important gap in understanding the interplay between the network structure and the neighbours’ performance in user innovation. More broadly, these results suggested that the interplay between the neighbours and the network structures merits attention that even goes beyond user innovation.

1 Graduate School of Arts and Sciences, The University of Tokyo, Tokyo, Japan. 2 Research Fellowship for Young Scientists (DC2), Japan Society for the Promotion of Science (JSPS), Tokyo, Japan. *email: yangkunhao@g.ecc.u-tokyo.ac.jp; ueda@gregorio.c.u-tokyo.ac.jp
**Introduction**

Any recent studies have shown that regular users, rather than specialists or organisations, have begun to capture a leading role in innovation (Gustafsson et al., 2012; Kratzer and Lettl, 2008; Lüthje, 2004; Reichwald et al., 2004; Von Hippel, 1986). This new perspective on innovation is called *customer innovation* or *user innovation* (Gustafsson et al., 2012; Von Hippel, 1986). User innovation occurs when a non-professional group of people contributes to innovation in a professional field (Howe, 2008). Because of its high effectiveness and low cost, it is considered a key approach to innovation (Hallikainen et al., 2019).

One important question about user innovation revolves around what features of an agent (person) affect her/his innovation ability, and how. Previous studies (Grosser et al., 2017; Perry-Smith and Mannucci, 2017; Perry-Smith and Shalley, 2003; Phelps et al., 2012; Shah et al., 2018) showed that two features have a significant impact: (1) the social network structure around the focal agent and (2) innovation abilities of other agents who have direct interactions with the focal agent (hereafter, we call the focal agent ego and these other agents neighbours). Researchers found that an open-network structure, which comprises neighbours in groups without ties between one another (See Fig. 1), benefited the innovation capability of an ego as the open-network structure provides a diverse information set to the ego (Perry-Smith and Mannucci, 2017; Phelps et al., 2012; Shah et al., 2018). In addition, researchers showed that neighbours with a high innovation performance also benefited the innovation of an ego as they provided valuable experiences and resources (Grosser et al., 2017; Shah et al., 2018). However, all these studies discussed the network structure separately from the innovation capability of the neighbours. In contrast, some theorists in network science (Lin, 2002; Pentland, 2015) suggested that there should be an interplay between these two factors. In other words, the effects of the network structure and of the neighbours will affect each other, dependently.

In this respect, on one hand, the direction of neighbours’ effect will differ depending on the openness of the network structure. The open-network structure will benefit the ego’s innovation because it brings heterogeneous neighbours to the ego (Newman and Dale, 2007; Prell et al., 2009; Rogers, 2010). These neighbours are in the various sub-groups within the network. Therefore, they can provide the nonredundant information for the ego (Burt, 2009; Capaldo, 2007; Kratzer and Lettl, 2008; Perry-Smith and Mannucci, 2017). Furthermore, the different knowledge backgrounds and behavioural patterns of these neighbours become the source of the nonredundant information for the ego (Newman and Dale, 2007; Prell et al., 2009; Rogers, 2010). The diversity and newness of nonredundant information will become the important resource for the ego’s innovation. By contrast, the enclosed network structure tends to undermine the ego’s innovation because it results in highly homogeneous neighbours (Newman and Dale, 2007; Prell et al., 2009; Rogers, 2010). Therefore, they can only provide the highly redundant information for the ego (Burt, 2009; Granovetter, 1983). Moreover, the enclosed network causes the so-called echo chamber phenomenon, which often happens among highly homogeneous neighbours (e.g., Pentland, 2015; Colleoni et al., 2014; Trevisanu and Hockema, 2009). Under this phenomenon, egos behave in the same way as their neighbours, resulting in a significant decline in innovation performance.

On the other hand, the intensity of the network structure’s influence will vary depending on the neighbours’ performance. First, the innovation performance of neighbours indicates the quantity of the (redundant or nonredundant) information resource neighbours can provide (Grosser et al., 2017; Shah et al., 2018; Lin, 2002). High-performance neighbours significantly affect the ego because she/he receives most of the information resources from them. In contrast, low-performance neighbours can hardly provide enough information resources to influence the ego. Subsequently, the ego will judge the underlying quality of this information when she/he is affected by her/his neighbours’ information (De Clercq and Dimov, 2008; Dimov and Milanov, 2010; Geringer, 1988; Podolny, 2001). The information from high-status neighbours will be evaluated more highly than that from low-status. Accordingly, it will affect the ego’s innovation more strongly. In addition, it is considered that, in a user innovation process, a neighbour’s status is decided by her/his previous innovation performance (Lin, 1999; Lin, 2002). Therefore, neighbours with high performance will be considered as highly creditable, and their information will have a larger impact on the ego’s innovation. In contrast, the information from neighbours without high performance will not be highly evaluated by the ego and will only have a small effect on the ego’s innovation (Perry-Smith and Shalley, 2003; Lin, 1999; Lin, 2002).

We combined the above two aspects to build the complete theoretical prediction about the interplay of the network structure and neighbours’ innovation performance: the network structure moderates the direction of the effect, and the neighbours’ innovation performance adjusts the intensity of the effect (as shown in Fig. 2).

As shown in Fig. 2, the information from high-performance neighbours will have a large effect on the ego’s innovation; if the network structure is open, the effect will be positive; if the network structure is enclosed, it will be negative. On the contrary, the information from low-performance neighbours will have less impact on the ego’s innovation, and the direction is decided by the network structure, in the same way as above.

Based on this theoretical prediction, we hypothesised as follows: in an open-network structure, neighbours with a higher innovation performance have a larger positive effect on an ego’s innovation; contrastingly, in an enclosed network structure, neighbours’ higher performance will have a larger negative effect on an ego’s innovation (hereafter denoted as the research hypothesis).

In the remainder of the paper, we conducted a panel regression to test whether and how the interaction between the network structure and innovation performance affect the ego’s innovation.

---

**Fig. 1 The illustration of the open-network structure and the enclosed network structure.** Both a and b show network structures of egos with the ego as a blue node, neighbours as yellow nodes and other nodes (users) having ties with neighbours as grey nodes. In a typical open-network structure (a), neighbours (in yellow) have no tie between one another, therefore, no neighbour can be approached through other neighbours of the ego, positioning the ego as a bridge connecting several (three in a) unconnected sub-groups in which her/his neighbours are. In a typical enclosed network structure (b), neighbours (in yellow), conversely, have ties between one another. As a result, every neighbour in (b) can be approached through other neighbours, indicating that the ego is stuck in one dense sub-group and has no chance of linking with other sub-groups in the network.
structure and neighbours’ innovation performance affects the innovation ability of each ego.

Additionally, our hypothesis was also tested by computer simulation in the supplement information (See SI 5). The regression model verified the moderating effect of the network structure on individual-level. However, the simulation further verified the moderating effect of the network structure in the network-level. However, since the results of these two ways were highly consistent, we only introduced the regression in the main text.

Methods

Raw data. It is, in general, difficult to study the research hypothesis due to the limited accessibility of data documenting both interactions between users and their innovation abilities. This study solved the problem by gathering data sets from a website called “Idea Storm”, which was designed by Dell to collect interesting ideas from their users. Idea Storm has records both on the idea submissions of each user and on interactions between users. This website is considered a reasonable target to observe user innovation (Bayus, 2013).

This research included neither ethical approval nor informed consent since “Idea Storm” is a public website and all users were anonymous.

We gathered two raw data sets from Idea Storm through a website crawler (See details about data collection in SI 1). One data set included interactions between users (denoted user network data hereafter) and the other documented ideas submitted by users (denoted idea submission data hereafter). A directed weighted network with 6,333 users was constructed using user network data; ties between users were defined by “vote” and “comment”. “Vote” represents a vote by one user to support another’s idea and “comment” represents a comment made by one user on someone else’s idea. We assumed that in the network, there was a tie between user A and B if either “vote” or “comment” existed between them (See Fig. 3). The direction and frequency of “vote” and “comment” were defined as the direction and weight of the ties.

Here, we did not distinguish the “vote” ties from the “comment” ties because the previous research (Bayus, 2013; Huang et al., 2014) using the same data from the “Idea Storm” suggested that both the “comment” and the “vote” have the same essence: the signal of preference for information and interaction. We also confirmed this argument using our data: We found that nearly all of the users (1,000 in 1,057) who sent a comment to another user also voted for the same user. This high synchronicity between the “comment” and the “vote” implies their essential similarity.

In the idea submission data, 326 users among 6,333 submitted 493 ideas defined as effective as Dell from 2007 to 2018. Based on Bayus (2013), we defined the effectiveness of ideas as the status of the ideas provided by Dell; when an idea was implemented or partly implemented into an actual product, it was regarded as effective.

Both the user network data and the idea submission data included timestamps showing when the interaction or the idea submission was conducted.

Panel data construction. We constructed an unbalanced daily panel data set at the user level based on the timestamps in the user network data and the ideas submission data. By this way, we got a data set which documented the change of the users’ antecedent network structure before they submitted every effective idea. Using this data, the regression model can distinguish the case where the dependent variable changed alongside with the independent variables from the case where the independent variables just changed after the change of the dependent variable.

To be more specific, the user network data documented the sender, the receiver and the timestamp for every comment or vote (See SI Fig. S2). The idea submission data recorded the effectiveness (namely, if this idea is effective or not) and the timestamp of every submitted idea.

To construct the network at a certain time point, we used all the comments and votes in the user network data before this time point. Therefore, the network can be considered as a result of the antecedent interactions between users (See SI Fig. S2). Then, the variables related to the network were computed based on the network at every time point. Other variables unrelated to the network were computed as the aggregation before the time point. Such as the effective idea submission at ti (introduced as the dependent variable in the following section) represented how many effective ideas had been submitted by the user before the time point ti. It is, thus, a cumulative variable.

By this way, we got the panel data set. The specific meaning of every variable in our data set was introduced in the “Variables for regressions” part.

Variables for regressions. Based on the raw data, we computed the variables for the panel regression models. The regression models tested whether highly innovative neighbours have a significantly positive (negative) impact on the innovation ability of an ego in an open (enclosed) network structure.

We set the innovation ability of an ego as the dependent variable. It was measured through a count variable, reflecting how many effective ideas an ego submitted into “Idea Storm”.

The openness of the network structure of an ego, one independent variable, was measured by constraint based on previous studies (Ahuja, 2000; Burt, 2009; Scott, 2017; Walker et al., 1997). The constraint, varying from 0 to 1, measures how many of the neighbours of an ego had ties with other neighbours. Since the ties in our data represent the preference allocation, this indicator means to what degree the ego and her/his neighbours shares the same preference on each other. The more open-
network structure indicates that an ego preferred the more diverse neighbours and information; while the more enclosed network structure indicates that an ego preferred the more homogeneous neighbours and information (Bayus, 2013; Cho and Shih, 2011). A constraint with the value of 1 indicates that every neighbour of an ego had ties with other neighbours directly or indirectly, meaning that the ego is embedded in one enclosed sub-group with members preferring to interact with each other instead of linking different sub-groups (as shown in Fig. 1b); a constraint with a value of <1 and >0 suggests that a higher proportion of neighbours had ties with other neighbours; and a constraint with the value of 0 indicates that none of the neighbours of the ego had ties with other neighbours, which means that the ego links, as a bridge in the network, several different preference-groups with one another. In summary, a smaller constraint indicates a more open-network structure for the ego (See details of calculation of constraint in SI 3).

The neighbours’ innovation performance, the other independent variable, was measured as the average number of effective ideas submitted by ego’s neighbours, because the total number of effective ideas has a fairly high-negative correlation with the constraint \( r = -0.43, p\text{-value} < 0.01 \). In order to consider the interaction between the neighbours’ innovation performance and the constraints in the model, a low correlation between these two independent variables is necessary to prevent multicollinearity. The average number of effective ideas for neighbours, although still significant, showed a low correlation with the constraint \( r = 0.08, p\text{-value} < 0.01 \) (Additionally, because of this, the number of neighbours each ego has was added as a control variable as explained below).

Moreover, to test our hypothesis, we examined the interaction between the constraint and the average number of effective ideas from neighbours.

In addition to the dependent and independent variables, we used four control variables: (1) the number of days from the ego has been observed in the data, (2) the frequency of interactions sent by the ego, (3) the number of neighbours of the ego, and (4) the number of all ideas (including the effective ideas and the ineffective ideas) submitted by the user. The number of observation days, the frequency of interactions, and the number of all idea submissions were added since it is natural to assume that when she/he spends more time, has more interactions and submits more ideas in the “Idea Storm”, an ego tends to be more likely to submit more effective ideas. “The number of neighbours of an ego was added because of the usage of the average number of effective ideas submitted by neighbours. Note that individual features were not included in the control variables. Therefore, to control the potential differences between individuals, we added the individual-effect into the panel model. As a result, the panel model will estimate the intercept separately for every user in the data. The potential individual differences were controlled by the intercept. This was explained in details in the following section.

In summary, the model is as follows:

\[
y_{it} = \alpha + \beta \cdot \text{idea submission} + \gamma \cdot \text{Constraint} + \delta \cdot \text{Frequency of interactions} + \epsilon \cdot \text{Number of neighbours} + \zeta \cdot \text{Number of all idea submission}
\]

The basic statistics including the number of data points, mean, standard deviation, minimum and maximum of each variable are shown in Table 1, and correlations and distributions between each pair of variables are shown in Fig. 4.

**Estimation approach.** Since the dependent variable was a count variable, we used the panel Poisson model with individual-effect to estimate the relationship between the dependent variable (the number of effective idea submissions) and the independent variables. Based on Hausman type test (Allison, 2005), the fixed-effects model was preferred (for the model without interaction, chi-square = 53.861, p-value < 0.01; for the model with interaction, chi-square = 18.738, p-value = 0.009). This model estimated the number of effective idea submissions by maximum likelihood estimation as follows:

\[
y_{it} = \lambda_{it} \cdot \text{idea submission of user } i \text{ at time point } t,
\]

\[
y_{it} = \Lambda(\alpha + \beta x_{it} + \gamma z_i)
\]

where \( \lambda_{it} \) is considered as a random variable with a Poisson distribution with a mean of \( \Lambda(\cdot) \).

---

**Table 1** Descriptive statistics of the panel data.

| Variable name                                      | Number of data points | Mean   | St. dev. | Min.   | Max.   |
|---------------------------------------------------|-----------------------|--------|----------|--------|--------|
| The number of effective idea submission           | 2,213,782             | 0.078  | 0.701    | 0      | 48     |
| Days being observed                               | 2,213,782             | 1,622.453 | 1,381.535 | 1 (-1.091) | 1 (1.455) |
| Neighbours’ average daily idea submission         | 2,213,782             | 0.343  | 0.394    | 0 (-0.349) | 1 (10.257) |
| Number of neighbours                              | 2,213,782             | 11.658 | 31.704   | 1 (-0.597) | 365 (31.517) |
| The number of all idea submissions                | 2,213,782             | 0.087  | 0.756    | 0 (-0.115) | 53 (69.989) |

*The minimum and maximum after standardised*
Here, $x_{it}$ represents all of our independent variables and control variables that vary over time and individuals. All variables were standardised (representing that their means were adjusted to zero and standard deviations were adjusted to one) in the regression models.

Because of the anonymization in “Idea Storm”, our model did not have $z_i$, that is the individual feature, which does not vary over time. However, the $\alpha_i$ captured the unobserved differences between individuals. Therefore, this analysis controlled the influence of inaccessible individual features. In another word, in this research, we controlled the potential individual differences by the above-mentioned method instead of using the control variables. Also note that based on the convention in the previous research (Bayus, 2013; Cornwell and Trumbull, 1994), we did not report the values of the intercepts in the “Results” section, since every user (namely for every $i$) has a different $\alpha_i$.

**Results**

The results of the regression models are summarised in Table 2. The model without interaction between the constraint and neighbours’ innovation performance was built together with the model with the interaction. By comparing these two models, we can test whether the moderating impact of network structures (the research hypothesis) is statistically significant. Please note that all independent and control variables were standardised in all of the models.
In other words, the constraint may just correlate to the four average effective idea submission and the four control variables. It was just a positive impact in the model without interaction, whereas it had a significant negative impact (coefficient = −0.21 and p-value < 0.01) in the model with interaction. More importantly, in the model with interaction, both the constraint and interaction showed significant negative impacts (coefficient = −1.35 and coefficient = −1.89, both p-values < 0.01) on the dependent variable. This reflects that when an ego’s constraint is low, a larger average idea submission from neighbours leads to more effective idea submissions for an ego; conversely, when an ego’s constraint is high, a larger average idea submission from neighbours lowers the number of effective idea submissions by the ego. These results were also robust when we used the average moving time windows of 1 week and 1 month to construct the panel data (See SI 4). Thus, our results supported the research hypothesis that, in an open-network structure, neighbours with a higher innovation performance have a larger positive impact on an ego’s innovation; in an enclosed network structure, neighbours with a higher innovation performance have a larger negative impact on an ego’s innovation.

We visualised this result in Fig. 5:
The three lines in Fig. 5 represent three hypothetical users whose constraints are equal to the mean (the middle constraint), the mean plus one standard deviation (the high constraint), and the mean minus one standard deviation (the low constraint), respectively. All control variables of these three users are equal to the respective means. The x-axis denotes the average number of effective ideas submitted by their neighbours, and the y-axis denotes the logarithm of the effective ideas submitted by an ego, which was predicted by our panel Poisson model. Figure 5 showed that, under different intensities of constraint (high, middle, or low), the directions of the effect of neighbour’s performance changed. In an open-network structure (in case of low constraint), neighbours with larger average effective idea submissions have a larger positive effect on an ego’s effective idea submission; in contrast, in an enclosed network structure (in case of high constraint), neighbours with larger average effective idea submissions have a larger negative effect on an ego’s effective idea submission. The result was consistent with our research hypothesis.

However, there was still an alternative explanation for our results: the interaction between the two independent variables, constraint and neighbour’s average effective idea submissions, was just a by-product of the interaction between neighbours’ average effective idea submission and the four control variables. In other words, the constraint may just correlate to the four control variables. Therefore, the significant effect of the interaction of constraint and neighbours’ average effective idea submission may only indicate that the interactions of the four control variables and the neighbours’ average effective idea submission are significant. To rule out this alternative possibility, we built an additional model, including not only the interaction between the neighbours’ average effective idea submission and the constraint, but also the interactions between the neighbours’ average effective idea submission and all four control variables. The results of this additional model are shown in the Table 3.
In this model, the interaction of the neighbours’ average effective idea submission and the constraint was still significantly negative (coefficient = −0.14 and \( p\)-value = 0.05). This means that the negative interaction of neighbours’ average effective idea submission with constraint is not a by-product of the interactions of neighbours’ average effective idea submissions with other attributions (control variables) of the ego.

In addition, other significant interactions also implied that the direction of the effect of the neighbours’ average effective idea submission was affected both by the network structure and by the ego’s attributions. We consider that this result implies a direction of the future study and will discuss it further in the “Future studies”.

To summarise all aforementioned analyses, the results obtained robustly supported our research hypothesis.

Discussion

In this paper, we showed that there is a significant interplay between the openness of the network structures of the egos and the innovation performance of the neighbours. When the network structures around the egos are open, they experience a more positive effect on their innovation from their neighbours with higher innovation performance. Conversely, if the network structures around the egos are enclosed, their innovation experiences more negative effects from neighbours with higher innovation performance.

Overall implication. In previous studies, the innovation performance of neighbours and the network structure are always considered separately with no regard to the interaction between them (Capaldo, 2007; Dhanaraj and Parkhe, 2006; Grosser et al., 2017; Nooteboom, 2000; Nooteboom, 2006; Perry-Smith and Mannucci, 2017; Perry-Smith and Shalley, 2003; Shah et al., 2018). However, based on existing theories, the interaction is even more important (Granovetter, 1983; Lin, 2002). Indeed, the innovative neighbours are important for the ego’s innovation since they will strongly affect the ego. However, the effectiveness of the information provided by neighbours is moderated by the network structure (Burt, 2009; Lin, 2002). Only when considering the interplay between network structure and neighbours’ innovation abilities can we obtain a comprehensive understanding of the impacts of these two factors. Therefore, this research filled an important gap in understanding the interplay between network structure and neighbours’ performance in user innovation.

More broadly, a similar moderating effect from the network structure is also implied in other activities, such as online investment (Pentland, 2015). Our findings, thus, also point to more discussions about the interplay between network features beyond user innovation and indicate the limitations in solely considering a single network feature in future studies.

The implication of real-world application. We believe that our results also contribute to the findings in the management domain and organisation research: The social interactions do not always benefit the innovation; by contrast, sometime social interactions will even impair the potential innovation ability of the agent (Pentland and Feldman, 2007). Additionally, our results provided an important instruction for innovative organisation design. The interactions through an open-network structure will be beneficial, but those through an enclosed network may lead a contrary effect. Therefore, when managers try to design their organisation to be innovative, they should leave some “bridges” across different subgroups and departments so that the social network keeps open.

Future studies. For further discussions, we believe that our research can be expanded in the following four ways:

1. The additional model (in Table 3) indicated that the direction of the effect of neighbours’ innovation performance was affected both by the network structure and by the ego’s attributions. A “novice” ego (who spent only a short time in “Idea Storm” and had less neighbours and idea submissions) will be more negatively affected by the neighbours with high innovation performance in an enclosed network structure than an experienced ego (who spent longer time in “Idea Storm” and had more neighbours and idea submissions). This is believed to be because the high-status neighbours affect a “novice” ego more deeply than an experienced ego. However, this possibility requires further testing in future research.

2. Verifying our results within an information-flow network: The essence of the network in this research is the signal of preference and attention allocation. The ties between nodes showed who had a preference on whom. Although it is reasonable to consider the structure of the preference network and use the constraint to measure it (Cho and Shih, 2011; Vilhena et al., 2014), we expect that the effect of the network structure will become more evident when considering an information-flow network. In this type of network, the ties show information-flow relationship between nodes. Thus, nodes without ties between each other exchange no information. We expect that this stricter definition on the network would facilitate the effect of the network structure.

3. Merging with other data sets including individual features: Because of the anonymization in the “Idea Storm”, the individual features of users were inaccessible in our data set. Instead, we controlled the potential individual differences by the panel regression method. We did not choose other data sets with individual features because the “Idea Storm” documents both interactions between users and their innovation abilities. To the best of our knowledge, it is the only data set including both the network and the innovation performance in the user innovation domain. Of course, if there are other data sets, which can provide more individual information, some other interesting discussions between the relationship of individual features and the

Table 3 Additional panel poisson model for submission of effective ideas.

| Count variable: how many effective ideas were submitted by this user |
|---------------------------------------------------------------|
| Neighbours’ average effective idea submission | −0.94*** (0.06) |
| Constraint | −0.51*** (0.02) |
| Interaction of constraint and neighbours’ average effective idea submission | −0.14* (0.07) |
| The frequency of interactions | −0.08*** (0.01) |
| The number of days being observed | −0.01*** (0.00) |
| The number of neighbours | 0.25*** (0.01) |
| The number of all idea submission | 0.19*** (0.01) |
| Interaction of the frequency and neighbours’ average effective idea submission | −0.33*** (0.02) |
| Interaction of the number of days being observed and neighbours’ average effective idea submission | 0.04*** (0.01) |
| Interaction of the number of neighbours and neighbours’ average effective idea submission | 0.80*** (0.04) |
| Interaction of the number of all idea submission and neighbours’ average effective idea submission | 0.55*** (0.02) |
| The number of data points | 2,213,782 |
| Log likelihood | −129,375.98 |

*p < 0.1; **p < 0.05; ***p < 0.01 (standard error)
network features can be explored. Thus, in the future research, considerable effort should be allocated for data searching and assembling to enrich the "Idea Storm" by other data sets.

4. Analysing different interactive behaviours separately: in this research, because of the essential similarity between the "vote" and the "comment", we regarded these two behaviours as synonymous. However, it is also interesting to consider the effect of different networks shaped by the two different behaviours. Therefore, it is valuable for future studies to conduct this analysis, especially when their data includes different interactive behaviours.

Data availability
The data sets analysed during the current study are available in the Github repository: https://github.com/yoguluto/Palcomms.

References
Allison PD (2005) Fixed effects regression procedures for longitudinal data using SAS. SAS Institute, Cary
Ahuja G (2000) Collaboration networks, structural holes, and innovation: a longitudinal study. Adm Sci Q 45(3):425–455
Bayus BL (2013) Crowdsourcing new product ideas over time: an analysis of the Dell IdeaStorm community. Manage Sci 59(1):226–244
Burt RS (2009) Structural holes: The social structure of competition. Harvard University Press, Cambridge
Capaldo A (2007) Network structure and innovation: the leveraging of a dual network as a distinctive relational capability. Strateg Manag J 28(6):585–608
Cho T-S, Shih H-Y (2011) Patent citation network analysis of core and emerging technologies in Taiwan: 1997–2008. Scientometrics 89(3):795–811
Colleoni E, Rozza A, Arvidsson A (2014) Echo chamber or public sphere? Predicting political orientation and measuring political homophily in Twitter using big data. J Commun 64(2):317–332
Cornwell C, Trumbull VN (1994) Estimating the economic model of crime with panel data. Rev Econ Stat 76(2):360–366
Dhanaraj C, Parkhe A (2006) Orchestrating innovation networks. Acad Manag Rev 31(3):659–669
De Clercq D, Dimov D (2008) Internal knowledge development and external knowledge access in venture capital investment performance. J Manag Stud 45(3):585–612
Dimov D, Milanov H (2010) The interplay of need and opportunity in venture capital investment syndication. J Bus Venturin 25(4):331–348
Granovetter M (1983) The strength of weak ties: A network theory revisited. Socio Theory 1:201–233
Gerring JM (1988) Joint venture partner selection: strategies for developed countries. Quorum Books, New York, NY
Grosser TJ, Venkataramani V, Labianca CJ (2017) An alter-centric perspective on employee innovation: the importance of alters’ creative self-efficacy and network structure. J Appl Psychol 102(9):1360–1374
Gustafsson A, Kristensson P, Witell L (2012) Customer co-creation in service innovation: a matter of communication? J Serv Manag 23(3):311–327
Hallikainen H, Alamaki A, Laukkanen T (2019) Lead users of business mobile services. Int J Inf Manag 47:283–292
Howe J (2008) Crowdsourcing: how the power of the crowd is driving the future of business. Random House, New York, NY
Huang Y, Vir Singh P, Srinivasan K (2014) Crowdsourcing new product ideas under consumer learning. Manage Sci 60(9):2138–2159
Katzir J, Letef C (2008) A social network perspective of lead users and creativity: an empirical study among children. Creat Innov Manag 17(1):26–36
Lin N (1999) Social networks and status attainment. Annu Rev Sociol 25(1):467–487
Lin N (2002) Social capital: a theory of social structure and action. Cambridge University Press, Cambridge
Lüthje C (2004) Characteristics of innovating users in a consumer goods field: an empirical study of sport-related product consumers. Technovation 24(9):683–695
Newman L, Dale L (2007) Homophily and agency: creating effective sustainable development networks. Environ Dev Sustain 9(1):79–90
Nooteboom B (2000) Learning by interaction: absorptive capacity, cognitive distance and governance. J Manag Gov 4(1–2):69–92
Nooteboom B (2006) Innovation, learning and cluster dynamics. In: Asheim B, Cooke P, Martin R (eds) Clusters and regional development. Routledge, London, pp. 137–163
Pentland A (2015) Social physics: how social networks can make us smarter. Penguin, New York, NY
Pentland BT, Feldman MS (2007) Narrative networks: patterns of technology and organization. Organ Sci 18(5):781–795
Perry-Smith JE, Mannucci PV (2017) From creativity to innovation: the social network drivers of the four phases of the idea journey. Acad Manag Rev 42(1):53–79
Perry-Smith JE, Shalley CE (2003) The social side of creativity: a static and dynamic social network perspective. Acad Manag Rev 28(1):89–106
Phelps C, Heidl R, Wadhwa A (2012) Knowledge, networks, and knowledge networks: a research and review agenda. J Manag 38(4):1115–1166
Podolny JM (2001) Networks as the pipes and prisms of the market. Am J Socio 107(1):33–60
Prell C, Hubacek K, Reed M (2009) Stakeholder analysis and social network analysis in natural resource management. Soc Nat Resour 22(6):501–518
Reichwald R, Seifert S, Walcher D (2004) Customers as part of value webs: towards a framework for webbed customer innovation tools. In: Sprague RH (ed.), Proceedings of the 37th Annual Hawaii International Conference on System Sciences, Big Island, 5–8 Jan 2004
Rogers EM (2010) Diffusion of innovations. Simon and Schuster, New York, NY
Scott J (2017) Social network analysis. SAGE Publications, California
Shah NP, Levin DZ, Cross R (2018) Secondhand social capital: boundary spanning, secondhand closure, and individual performance. Soc Netw 52:18–27
Steverijnas M, J, Hockema S (2009) The value of the unpopular: counteracting the popularity echo-chamber on the Web, 2009 IEEE Toronto International Conference Science and Technology for Humanity (TIC-STH), Toronto, 26–27 Sept 2009
Vilhena DA, Foster JG, Rosvall M, West JD, Evans J, Bergstrom CT (2014) Finding cultural holes: how structure and culture diverge in networks of scholarly communication. Socio Sci 1:221–238
Von Hippel E (1986) Lead users: a source of novel product concepts. Manag Sci 32(7):781–805
Walker G, Kogut B, Shan W (1997) Social capital, structural holes and the formation of an industry network. Organ Sci 8(2):109–125

Acknowledgements
This study was supported by JSPS KAKENHI Grant Number JP16H01725 and MEXT KAKENHI Grant Number JP26118002 for the third author.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1057/s41599-019-0383-x.

Correspondence and requests for materials should be addressed to K.Y. or K.U.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.