The Effect of Organizational Structure on Open Innovation: A Quadratic Equation

Junyeong Lee\textsuperscript{a}, Jinyoung Min\textsuperscript{b,*,} Heeseok Lee\textsuperscript{a}

\textsuperscript{a} College of Business, KAIST, Seoul, Republic of Korea
\textsuperscript{b} College of Business, Chosun University, Gwangju, Republic of Korea

Abstract

Open innovation has recently begun to receive increasing attention in organizational research. Studies on open innovation have emphasized the impact of organizational structure on both inbound and outbound open innovation but have tended to focus on the former. This study attempts to provide a better understanding of open innovation within the context of open source by examining 2,811 projects, especially in GitHub. The analysis results reveal that the decentralization of decision-making encourages both inbound and outbound open innovation. The impact of the decentralization of decision-making differs between inbound and outbound open innovation, a finding that both scholars and practitioners must consider.

Keywords: Decision-making Structure; Decentralization; Open Innovation

1. Introduction

Open innovation has received much attention in organization research due to its numerous benefits, such as its sharing of risks and resources with partners and its better adaptation to the needs of the rapidly changing market [1, 2]. Firms are increasingly adopting open innovation strategies [3], the paradigm of which consists of two dimensions: inbound open innovation and outbound open innovation. The literature on open innovation has explored both open innovation types [4, 5]; however scholars have mainly focused on inbound open innovation [6]. Open innovation is becoming more popular, with numerous innovation platforms and external participant developments appearing on the Internet [7]. To put ideas from external sources into action, each organization must identify the best ideas, conduct internal feasibility and profitability analyses, and implement them in development projects. For this purpose, organizations should develop organizational architecture suitable for open innovation, including structures, processes, and routines [4, 8].

\* Corresponding author. Tel.: +82-62-230-6812.
E-mail address: saharamin@chosun.ac.kr.
Organizational structure is an important managerial lever for open innovation [9, 10]. Since the development of internal networks is required for the successful management of acquired external knowledge [11], this study concentrates on an important aspect of organizational structure: the degree of decentralization of decision making. Effective decision and organizational architecture can contribute greatly to innovation success [12, 13]. The research has repeatedly shown that organizational performance decreases when control is strengthened through an increased centralization of authority [13, 14]. The initiative of subordinates and innovative ideas may be discouraged under a centralized decision-making structure [15], as organizations face structural difficulties in generating innovation: “Along this tree from foot to crown, ideas flow up and vetoes down” ([16] p. 276). In other words, performance is positively associated with the decentralization of decision making. However, a centralized structure is also important for open innovation. For example, P&G uses centralized decision making in its Connect and Develop strategy since it fosters risk-taking, monitoring, and long-term thinking [3, 17]. Thus, the important question is how innovation is affected by the centralization or decentralization of the decision-making authority [18]. The decision making structure plays an important role in the success of open innovation.

We therefore investigate the role of decision-making structures in open innovation through hypotheses with quadratic terms. The objective of our analysis is to study the effect of the decision making structure on the degree of both inbound and outbound open innovation at the project level. Decision making usually occurs at the project level, as many companies develop their firm-level capabilities through innovation projects [1, 19]. Hence, project-level studies can provide a deeper understanding of decision-making processes in open innovation [19]. To answer our research question, we explore the open source context, specifically the social coding site GitHub, because the mechanisms of the development of open innovation offer a wide range of alternatives, including open source platforms [20].

2. Conceptual Background and Hypotheses

2.1. Open innovation (inbound/outbound)

Chesbrough, Vanhaverbeke and West [2] have suggested two conceptually distinct dimensions of open innovation: inbound, or outside-in, open innovation and outbound, or inside-out, open innovation. Since then, the literature on open innovation has emphasized both dimensions [4, 5]. However, past studies have concentrated more on inbound open innovation [21], although every inbound open innovation in one organization generates a reciprocal outbound open innovation from another [22]. The effectiveness of open innovation depends on the ability to handle decentralized innovation processes [23], and organizations can benefit from knowledge voluntarily provided by external participants [7, 24]. In this sense, the literature on Free/Libre Open Source Software (FLOSS) can offer insight into the open innovation concept by, for example, arguing that addressing diffuse needs is important for the success of FLOSS projects since this can attract contributors [25]. In sum, past studies of open innovation implies that the consideration of both inbound innovation and outbound innovation is important within the context of open source.

2.2. Organizational structure: Structure of decision making

Previous studies have investigated the design of decision-making structures from many perspectives, since their properties are multifaceted [26]. The literature on the decentralization of information has consistently revealed that organizational structures can be designed to optimize efficiency under limited capacity [27, 28]. By contrast, Sah and Stiglitz [29] and others [30, 31] have focused on human evaluation, identifying two kinds of decision-making structure and explaining how they differ systematically in terms of human fallibility. Dessein and Santos [32] developed a theory for integrating task specialization and coordination in the analysis of organizational structure. Their basic idea is that decentralizing decisions facilitates the usage of local knowledge
and offers local managers the flexibility to modify their actions, which should adapt to a changing environment. Organizations must allow a consideration of the full range of organizational structures in order to advance a general treatment of decision making [26].

There are two kinds of organizational decision-making error: omission error and commission error. An omission error is the rejection of a good project and is described as a “Type I error” in terms of statistics inference. A commission error is the acceptance of a bad project and is described as a “Type II error” [33]. Each type of decision-making structure reduces the probability of either omission or commission error. For example, the incidence of commission errors is relatively high and that of omission errors relatively low in decentralized decision making. The optimal choice of organizational structure significantly depends on a comparison between the two types of error [34]; by extension, these two error types can lead to exploration and exploitation [35].

Csaszar [35] borrowed from decision theory to analyze the relationships between the two error types in terms of organizational exploration and exploitation, as decision theory is based on losses while economics is based on gains (i.e., utility). Decision theory seeks to develop optimal decision rules under conditions of imperfect knowledge such as the state of nature, the actions of decision makers, and the consequences of those actions. Optimal decisions are made amid deviations, consisting of both omission errors and commission errors, since most decision-making tasks risk these two types of error [35, 36]. Overall, prior research of organizational structure indicates that decision making structure is critical in open innovation output.

2.3. Hypotheses

A high degree of decentralization indicates the high involvement of low-echelon decision makers with their many goals and values. Decentralized decision making increases the input of diverse goals and values into decision-making behavior and has a positive relationship with organizational performance [37]. Wu [18] strongly favored decentralized decision structures since they perform better based on the observation. Garicano and Posner [38] also argued that an organization should be decentralized to maximize the possibility that numerous fresh new ideas will be produced. The exploratory innovation of units is fostered by a decentralization of decision making [39], facilitating problem solving by widening the range of options for responses and supporting exploratory learning [40]. A decentralized decision-making structure also encourages widespread participation and fosters innovation from a broader variety of perspectives [41]. Thus, we hypothesize the following:

H1. As the decision-making structure becomes more decentralized, (a) the amount of inbound open innovation increases (b) the amount of outbound open innovation increases.

We established the first hypotheses based on some of the relevant literature. Other studies have produced contrary findings, however, arguing that centralized decision–making structures can increase open innovation depending on their screening processes and improvements in their innovation projects [42, 43]. A decentralized structure can facilitate communication and the awareness of potential innovations [44] and is effective in initiating innovation, including the widest-ranging exploration and idea generation [45, 46]. However, it is difficult to recognize or disseminate innovation under a structure that is too decentralized [47]. Organizations may change their structure or re-balance it after an initial period of decentralization, since it is difficult to achieve both exploration and stability simultaneously through a single organizational structure. The temporal sequencing of centralization and decentralization can lead to higher performance [48]. Following the literature, we suggest the diminishing effects on open innovation as decision-making becomes more decentralized. Thus, we hypothesize the following:

H2. A decentralized decision-making structure has a diminishing effect on (a) inbound open innovation (b) outbound open innovation: there exists an inverted U-shaped relationship between the decision-making structure and (a) the amount of inbound open innovation (b) the amount of outbound open innovation.
3. Method

3.1. Research Context (GitHub)

Open source is an open innovation that features the collaborative development of the project and confers shared rights to use the outcomes [49]. Our research question is tested in GitHub, which provides “social coding” tools to support the “git” version control systems for source code dispersal control management. The site offers a collaborative code review system that facilitates discussions among developers, reviews code suggestions, and compares projects’ progress [70]. GitHub currently hosts over 12.3 million code repositories and has 5.8 million developers. GitHub hosts many large multi-developer projects that have been ongoing for a long time. Each repository on GitHub has a devoted project page that hosts the source code files, commit history, pull requests, and other project data [50]. Many studies have based their research questions on GitHub (e.g. [50, 51]).

3.2. Data Collection

We tested our hypotheses on a dataset drawn from the GitHub website (http://GitHub.com). This raw data were gathered using a Web crawler developed by Python; it automatically downloads the target HTML pages and parses their content to extract the variables that contain issues, pull requests, collaborators, releases, commits, and other control variables. To avoid sampling problems, we collected the data based on categories suggested by Chen, Li, Clark and Dietrich [52]. We then eliminated missing, invalid, and redundant data; for instance, data on projects without issues, collaborators, or other control variables are eliminated. Our final sample comprises 2,811 data items.

3.3. Measures

3.3.1. Dependent Variables

We define inbound open innovation as the number of commits. Decision makers, as collaborators, can make commits (i.e., code revisions) by modifying project outcomes by accepting external suggestions. Commits are closely linked to relationships with external participants for the enhancement of internal innovation [9], as commits don't happen if decision makers don't take external suggestions. This mechanism allows us to describe commits as inbound open innovation.

We define outbound open innovation as the number of releases. Outbound open innovation is based on external relationships for the commercial use of knowledge [9]. Each project can release project output in order to distribute and share artifacts outside of the project. We thus consider the release of a project as outbound open innovation.

3.3.2. Independent Variables.

Christensen and Knudsen [26] defined two types of organizational structure. At one pole, low-echelon managers report to their immediate superiors in hierarchical organizations; the other extreme allows substantial variation in the number of people who can make independent, parallel decisions in decentralized structure. Following the definition of “architecture” in Augier and Knudsen [42] and the description in Christensen and Knudsen [26], we measure the degree of the decentralization of decision making as the number of people who can make independent and parallel decisions at the project level, since decision rights are the rights of individual members to make decision. In our research context, the collaborators, members with code revision rights, can merge suggested pull requests, and reflect on the causes of the issue. As the number of collaborators increases, the number of people with decision-making rights in their project also increases, producing a more decentralized
decision-making structure. Thus, we use the number of collaborators in a project to define the degree of the decentralization of decision-making.

3.3.3. Control variables

Our study controls several variables for project-related factors based on the literature.

We control for external suggestions, since external knowledge sources and the breadth of the external innovation linkages can lead to innovation outcomes and thus improve financial performance [53, 54]. To control suggestions, we use pull requests and issues suggested by external participants, who can provide suggestions and offer project knowledge by submitting an issue or a pull request, thus helping improve the artifacts of the project [50]. We calculate a factor score for the total number of pull requests and issues to represent external suggestions.

A previous study controlled the general interest level in the project, defined as the number of page views [55]. We use the number of “watch” and “star” notifications, which are similar to page views. We compute a factor score for these factors to reflect the general interest level. A previous study controlled the amount of project information, operationally defined as the project’s level of descriptive detail [56]. We use this definition to control for the effects of our dependent variables. We use the number of days as the project age in this study, as the numbers of commits and releases are likely to increase with the age of the project [55]. We use project category as a control variable to control the possible effect caused by differences among project categories [52].

4. Analysis and Results

We conducted a data transformation and verified potential collinearity problems for further analysis. We analyzed our hypothesized association in our research model using linear regression with the transformed data.

4.1. Data Analysis

Our data were skewed and non-negative, with many zero values, since both kurtosis and skewness are greater than 2 for all variables. To employ the data for further study and alleviate the potential heteroscedasticity problem associated with each variable, we normalized each variable by adding 1 and used it in logarithmic form (log10 Variable); normalizing variables after adding 1 is useful because taking logs without adding 1 can discard observations with non-positive values [57]. After the log transformation, all variables are found to be acceptable for proving normal distribution, since both the skewness and kurtosis of most variables (except the kurtosis of project age) are between -2 and +2 [58].

The average number of decision–making structure is 8.54. The average number of inbound open innovation is 525.60, and the average for outbound open innovation is 10.28. The average age is 965.79 days (around 2.646 years). Suggestions and general interest level are reflected by two actual variables. The average numbers of pull requests and issues (i.e., suggestions) are 35.47 and 50.91, respectively. The average numbers of watches and stars (i.e., general interest level) are 26.59 and 185.90, respectively.

After normalization, we analyzed the observed data using linear regression analysis through SPSS. We did not need to demonstrate measurements such as validity, reliability, or a common method bias, as these were observed data. Instead of a measurement test, we checked for possible collinearity problems to ensure the validity of the regression results [59]. The variance inflation factor (VIF) values of the independent variables were found to be less than the standard value of 10 (VIF<10), indicating that multicollinearity was not a serious concern [60].

4.2. Analysis results

Table 1 shows the descriptive statistics of the variables and the correlations between the log transformed variables. Suggestions and general interest level were assessed from the factor scores. Their mean value and
standard deviation are 0 and 1, respectively; the variables are not correlated. All relationships between variables are significant except for that between suggestions and general interest level.

Table 1. Descriptive Statistics and Correlation Matrix of Transformed Data

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
<th>(9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Decentralization of decision making</td>
<td>0.690</td>
<td>0.407</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2) Inbound open innovation</td>
<td>1.957</td>
<td>0.768</td>
<td>.703**</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3) Outbound open innovation</td>
<td>0.449</td>
<td>0.585</td>
<td>.531**</td>
<td>.599**</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(4) Suggestions</td>
<td>0.00</td>
<td>1.00</td>
<td>.723**</td>
<td>.660**</td>
<td>.505**</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(5) General interest level</td>
<td>0.00</td>
<td>1.00</td>
<td>.350**</td>
<td>.258**</td>
<td>.222**</td>
<td>0</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(6) Project information availability</td>
<td>3.145</td>
<td>0.607</td>
<td>.213**</td>
<td>.212**</td>
<td>.223**</td>
<td>.143**</td>
<td>.345**</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>(7) Project age</td>
<td>2.851</td>
<td>0.380</td>
<td>.276**</td>
<td>.339**</td>
<td>.280**</td>
<td>.173**</td>
<td>.126**</td>
<td>.086**</td>
<td>1</td>
</tr>
</tbody>
</table>

** Correlation is significant at the 0.01 level (two-tailed)

Table 2 shows the overall test results. All coefficients are standardized. First, we analyze our model step by step in order to evaluate its suitability. For each dependent variable, model 1 shows a base model that includes only the control variables. Model 2 includes the direct effect of the decentralization of decision-making, and model 3 includes the square form of the decentralization, our research model. Based on the residual sum of the square, sample size, and the number of predictors, we calculate the relative suitability of the models (i.e., between models 1 and 2 and models 2 and 3). The F-test values show that our research model is well-established.

After checking the goodness-of-fit of our research model, we test the hypotheses. The results show that the degrees of both inbound and outbound open innovation increase as the decision-making structure becomes more decentralized. As predicted, the decentralized structure encourages both inbound and outbound open innovation; hypotheses 1a and 1b are thus supported.

Based on the effect the quadratic form of decentralized decision-making has on each open innovation type, we check our second hypotheses regarding the tendency of the decision-making structure. The analysis results show that the coefficient of the quadratic form of the decision-making structure’s effect on inbound open innovation is negative, indicating that the effect diminishes as decision-making becomes more decentralized and that there exists an inverted U-shaped (i.e., concave) relationship between decision-making structure and the degree of inbound open innovation. Thus, hypothesis 2a is supported. By contrast, the analysis results show that the coefficient of the quadratic effect of the decision-making structure’s on inbound open innovation is positive; thus, hypothesis 2b is not supported. This unexpected result indicates that the decision-making structure’s effect on outbound open innovation grows as the decision-making becomes more decentralized and that there exists a convex relationship between decision-making structure and the degree of outbound open innovation.

The results show that all control variables are significant for both open innovation types. The degrees of both inbound and outbound open innovation increase as project age increases. Project information availability and general interest levels also have significant influences on open innovation, since those project-related factors can signal the attractiveness of participation to external participants [55, 56]. Suggestions are a critical factor for both open innovation types, since its coefficient is the greatest among control variables.
To confirm the results of our study, we perform a robustness check given the potential heteroscedasticity problem. We employ the heteroscedasticity robust standard error [61] to examine the potential heteroscedasticity issue instead of homoscedasticity in an OLS regression. In the model with heteroscedasticity, all relationships are maintained.

Table 2. Linear Regression Results

<table>
<thead>
<tr>
<th>Independent variable</th>
<th>Inbound open innovation</th>
<th>Outbound open innovation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Categories (11 dummy variables)</td>
<td>Included</td>
<td>Included</td>
</tr>
<tr>
<td>Age</td>
<td>.199***</td>
<td>.163***</td>
</tr>
<tr>
<td>Project information availability</td>
<td>.017**</td>
<td>.041**</td>
</tr>
<tr>
<td>General interest level</td>
<td>.218***</td>
<td>.109***</td>
</tr>
<tr>
<td>Suggestions</td>
<td>.614***</td>
<td>.378***</td>
</tr>
<tr>
<td>Decentralization of decision making</td>
<td>.338***</td>
<td>.380***</td>
</tr>
<tr>
<td>Quadratic form of Decentralization of decision making</td>
<td>- .063***</td>
<td>.052*</td>
</tr>
<tr>
<td>R2</td>
<td>0.552</td>
<td>0.591</td>
</tr>
<tr>
<td>Adjusted R2</td>
<td>0.55</td>
<td>0.588</td>
</tr>
<tr>
<td>AIC</td>
<td>-3714.27</td>
<td>-3963.92</td>
</tr>
<tr>
<td>F-test between model</td>
<td>261.67***</td>
<td>15.54***</td>
</tr>
</tbody>
</table>

All variables are common-log transformed.
* p < 0.05; ** p < 0.01; *** p < 0.001

5. Discussion

Our study confirms the effect of decision-making structure on open innovation using data from 2,811 open source projects. The analysis results show that, first, both inbound and outbound open innovation increase as the decision-making structure becomes more decentralized. A decentralized structure reduces the probability of omission errors (the rejection of good suggestions) and increases the probability of commission errors (the acceptance of bad suggestions) [34]. Despite this tradeoff, decentralized decision-making is found to enhance both inbound and outbound open innovation, showing that omission errors are more critical for open innovation than commission errors because open innovation is still encouraged even though commission errors may occur under decentralized structure. The finding also implies that exploration, which is related to omission errors, is more important in open innovation than exploitation is.

Our result shows the different tendencies of inbound and outbound open innovation in their relationships with decision-making structures. The concave relationship between decision-making and inbound open innovation indicates that the effect of the relationship decreases as decision-making becomes more decentralized, as hypothesized in our study. This result is consistent with the controversial findings of previous studies (e.g.,[48]) and examples (e.g., the “heart-bleed” bug). However, the convex relationship between decision-making structure and outbound open innovation indicates that the effect of the relationship increases as decision-making becomes
more decentralized. A plausible reason for this interesting result can be drawn from bounded capacity [62]. From within its focal domain, each project may inhibit related innovation activities. The bounded capacity of the focal domain limits any increase in inbound open innovation. However, there is no limitation on the bounded capacity in outbound open innovation since it proceeds outside of the focal project and can be applied to various other projects. Another possible explanation is that many project organizations provide only little knowledge but use much external knowledge [8]. There are many vacancies to fill and much knowledge to expand in outbound open innovation.

5.1. Future research opportunities

Several future research options will be of interest. First, we collected our data by crawling through the GitHub website. The data is real but represents only a snapshot of the dynamics of open innovation in open source projects. They may not reflect all of the dynamics of open innovation. To overcome these data limitations, future research should collect longitudinal or more comprehensive data such as conversation logs, which will provide richer study material and allow a deeper examination of open innovation. Second, we collected data from one source, GitHub, and the sample of this study consisted of over 2,800 projects. In order to generalize the findings of this study, future research should include other research contexts. In addition, Watts [63] suggests that a decentralized structure can sometimes be arranged, as it has the ability to develop collective intelligence. Participants ideally contribute to better project outcomes in modular ways and help form the conditions of the activity and thus help build projects with a decentralized structure rather than hierarchically assigned tasks [64]. In line with these findings, future research should consider the structure of decision-making with collective intelligence and modular tasks in open innovation.

5.2. Theoretical implications

We explored the importance of decentralized decision-making in open innovation. Following the concept proposed by Sah and Stiglitz [29], many studies have pointed out the importance of decision making structure; however, they have usually adopted economical modeling or simulation (e.g., [26, 48]). Few studies have investigated the effect of structure empirically (e.g., [33]). Using real field data on the open source process, our study empirically examines the effect of the decision-making structure on open innovation. The results open the possibility of further empirical exploration.

Most studies focus on the effect of inbound open innovation on a firm; however, our study explores the relationships with both inbound and outbound open innovation at the project level. Using the project as the unit of analysis provides a more natural perspective since decisions are usually made at the project level [19].

Our study shows the contrasting effects of decision-making structure types on inbound and outbound open innovation. Inbound open innovation has an inverted-U relationship, while the convex effect of the decentralized structure on outbound open innovation is contrary to our hypotheses. These results hint the similar and different roles of the decision-making structure on both inbound and outbound open innovation. This unexpected finding will provide opportunities to further investigate the relationships between decision-making structures and open innovation.

5.3. Practical implications

Our findings reveal that the decentralized decision-making structure encourages both inbound and outbound open innovation. The decentralized structure provides new perspectives from outside the project that are not available inside. It also facilitates the identification of good proposals and leads to better project performance. A decentralized workflow nurtures discussion before the acceptance or rejection of suggestions from external
participants. Decision makers can handle the issues during code reviews such as conflict resolution based on individual testing and the continuous integration environment [51]. Decentralization also encourages the participation of external actors who want to show their ability to improve the project outcome. Decentralization can disperse labor-intensive tasks such as Q&A and debugging as individualized tasks. Our study suggests that practitioners should consider the decentralized decision-making structure as a way of unleashing open innovation.

References


