# The role of lead users in knowledge sharing

Lars Bo Jeppesen [a,*], Keld Laursen [b,1]

[a] Department of Innovation and Organizational Economics, Copenhagen Business School, Kilevej 14A, 2000 Frederiksberg, Denmark
[b] DRUID, Department of Innovation and Organizational Economics, Copenhagen Business School, Kilevej 14A, 2000 Frederiksberg, Denmark

## ARTICLE INFO

## ABSTRACT

This paper introduces a model of knowledge sharing in an online community of practice that suggests that knowledge contributions will be made by those who possess the relevant knowledge. For them, matching a ready-made solution to a problem is low cost. We hypothesize that lead users – due to their characteristics – are likely to possess more relevant solution knowledge and thus be centrally involved in contributing knowledge. Our results support the hypothesis by showing that lead user characteristics relate positively to making contributions to the community. In addition, we find that search and integration of knowledge from different external sources of relevance to the community positively moderates knowledge contributions by lead users.

## 1. Introduction

This paper focuses on the role of members' practice-related characteristics in explaining contributions to online communities of practice. We define contributions as voluntary acts of helping others by sharing knowledge. Consider an online community in which the members' interests and discussions revolve around certain practices such as the use and development of specific tools or a product. It is likely that members will be confronted with certain problems, and that some members have already found solutions to these problems. Prior literature within this area of interest focuses mostly on member motivation as being central to contributions to online communities of practice (see for instance, Constant et al., 1996; Kollock, 1999; Wasko and Faraj, 2000; Wasko and Faraj, 2005). While this literature has been tremendously valuable in establishing our understanding of the phenomena, we suggest adding a further aspect that can help explaining the sources of contributions.

In this paper, we center on the characteristics of individuals who are capable of matching up their solutions to the problems presented in the community. We pay particular attention to prominent features of information technologies used in online communities of practice that allow these capable individuals to solve problems. In line with recent research on contribution behavior in online communities (Ma and Agarwal, 2007; Olivera et al., 2008), our argu-ments are based on the view that information technologies give rise to differential modes of communication that can either inhibit or facilitate knowledge contributions. In the online community setting, a central mode of communication is "posting", which makes problem information widely visible and allow members with a solution to self-select to become solution providers at a low cost because a solution is already within reach. A second important mode of communication of online communities is that discussions typically concern novel problems that are "ahead of the archive." Thus, in order to be able to provide a solution, an updated knowledge base is necessary.

We argue that in online communities of practice, contributions are likely to stem from those individuals that hold the updated knowledge related to the practice in question. Primarily, these individuals are likely to be able to match-up and provide answers to questions for help. Our main hypothesis is therefore that a certain type of contributors, known as lead users (von Hippel, 1986), have a high propensity to contribute knowledge to online communities of practice. Lead users of a given product or service type have three defining characteristics (von Hippel, 1986; Morrison et al., 2000) that make them key agents in the knowledge sharing process: first, they are early adopters of the product or service; second, they experience the need for a given innovation earlier than the majority of the target market; and third, they are users who expect attractive innovation-related benefits from a solution to a problem.

We test our central hypotheses related to knowledge sharing empirically using data on knowledge sharing in an online community of practice focused on the development and improvement of computer-controlled music instruments in a public forum on the Internet. To obtain the relevant data, we carried out a survey, collected web log information and conducted interviews.

* Corresponding author. Tel.: +45 3815 2948; fax: +45 3815 2540.
E-mail addresses: lbj.ino@cbs.dk (L.B. Jeppesen), kl.ino@cbs.dk (K. Laursen).
[1] Tel.: +45 3815 2565; fax: +45 3815 2540.

## 2. Rationale

There are several reasons for studying lead users' knowledge sharing behavior in online communities of practice. First, there is growing evidence that decentralized and distributed knowledge communities are becoming an important factor in the economy. For instance, community based innovation has been found to play a key role in the creation of open source software products that have gained significant market shares from state-of-the-art commercial software (Lerner and Tirole, 2002; Franke and Shah, 2003). Furthermore, the sheer number of online communities has grown significantly in recent years (Ma and Agarwal, 2007).

Second, knowledge sharing is a central issue of relevance for individuals (Borgatti and Cross, 2003), teams and groups (Orr, 1996; Edmondson, 1999). It is seen as having an important impact on knowledge creation (Nonaka, 1994; Tsai and Ghoshal, 1998; von Krogh, 1998; Hansen, 1999) and as being an indicator of organizational capability and organizational performance (Kogut and Zander, 1992). Information technologies facilitate knowledge sharing and the enablement of communal resources (von Krogh, 2002).

Third, it is well established that lead users are an important source of value creation in the innovation process. As early adopters with strong needs for a solution, lead users become the sources of new product concepts when they go ahead and create novel solutions where no commercial solutions are available (von Hippel, 1986; Morrison et al., 2000). The importance of studying the knowledge sharing behavior of lead users is evident as such users – when connected – may provide needed know-how and task information; thus sustaining the level of knowledge supply in the community.

## 3. Theory

### 3.1. Online communities of practice and knowledge contribution behaviors

Over the last decade, new institutions and organizational forms have emerged in the form of knowledge intensive communities of practice in which knowledge is created and diffused among the community members and beyond. In this context, knowledge sharing is important as it sustains learning, but also because it supports processes of technological innovation, adoption and usage (Lave and Wenger, 1991; Kogut and Zander, 1992; Tsai and Ghoshal, 1998; Brown and Duguid, 2000; Gallivan, 2000). The vehicles of knowledge sharing include face-to-face interaction in geographically bound communities and, increasingly, electronic discussion groups (Constant et al., 1996), and Internet-based community interaction of distributed and loosely connected collaborating individuals (Armstrong and Hagel, 1996; Goodman and Darr, 1998; Muniz and O'Guinn, 2001; Benkler, 2004).

Earlier research on the use of information technologies to share knowledge confirms that employees in some geographically separate organizational divisions do share knowledge and help others, including organizationally remote strangers they will never meet in person (Finholt and Sproull, 1990; Constant et al., 1996; Lakhani and von Hippel, 2003). Our focus is on helping behavior in settings of online communities in the context of a public online community of product users that collaboratively discuss and develop modifications and new product concepts of a commercial product. Individuals involved share a common interest in a product and its tools but do not have direct ties or interdependencies, and they are not employed by the product owner organization and are not subject to direct pay, formal promotion and evaluation systems. Previous literature that has dealt with the opportunities and challenges raised by the emergence of these and similar communities has been valuable to our understanding of how they may add value to the processes and products of the communities themselves (Constant et al., 1996; Kollock, 1999; Raymond, 1999; Moon and Sproull, 2000; von Hippel and von Krogh, 2003) and to established firms (Jeppesen and Frederiksen, 2006) in terms of concrete products and innovations.

Prior research on contributions to electronic communities of practice has also taken interest in why individual motivation should lead to contributions. A range of studies has investigated motivation for knowledge sharing focusing on the dimensions of intrinsic and extrinsic motivations (Lerner and Tirole, 2002; Hertel et al., 2003; Lakhani and Wolf, 2005). That both of these dimensions jointly may affect likelihood of contribution have been suggested in earlier literature: a study by Wasko and Faraj (2000) suggested that individuals are motivated intrinsically to contribute knowledge to others because intellectual pursuits are challenging and fun. A later study by Wasko and Faraj (2005) studied the relationship between individual motivations and expertise and found that sharing one's knowledge is a function of members' position in the network, individual motivations such as the quest for reputation.

Our account adds a novel dimension to the discussion of contributions that can help us predict the sources of contributions. We submit that although motivation to share are a necessary condition for knowledge contribution, actually holding the necessary knowledge to make a sensible contribution will also increase the likelihood that a given individual go ahead and make such a contribution. Little is known about this knowledge-based dimension of contributions to online communities and, further, how contextual factors of the community may shape contribution and hence under which conditions we can expect contributions to be made.

Online communities of practice exhibit two prominent modes of communication of relevance for the knowledge sharing process. The first is "problem posting": we define posting as the disclosure of problem information to a pool of potential knowledge providers with a request for a solution/answer. In this context, studies of knowledge sharing in distributed environments with real problems to be solved have highlighted that under the condition of having problems posted to a crowd of potential solvers (knowledge providers), those who answer are those who happen to match-up their knowledge to the problem being presented. Problem posting creates awareness among many users, and given that an individual holds the relevant knowledge to the problem in question, the cost of providing an answer to this question becomes low and possibly explains why contributions are forthcoming in the first place. Prior research has suggested that those who do provide help in electronic networks are those that can do so at low cost because a solution is already within reach (Constant et al., 1996; Lakhani and von Hippel, 2003). In other words, members with a solution at hand self-select to become solution providers because they can essentially re-use existing knowledge in their contribution and avoid engaging in "from scratch" solution generation. One general feature of knowledge re-use is that agents invest in knowledge development and codification once, whereas the knowledge asset can be used repeatedly (Markus, 2001; Majchrzak et al., 2004; Haefliger et al., 2008). In this framework, the determinants of a contribution become the ability of any one potential provider to match a local solution to the presented problem. Due to the economics of re-use, the better the match, the lower the cost of answering and, other things being equal, the higher the likelihood of a contribution being made. In this view, having the right knowledge at hand is not only useful for solving the problem—it also generates the match which triggers the decision to provide knowledge.

The second prominent mode of communication in online communities of practice is that mainly problems that are "ahead of the archive" in terms of information content are posted in the community. The pertinent knowledge sharing process is characterized by the disclosure of "up-to-date problems" and the prevention of

exchanges of already solved problems. The technologies used in online communities of practice are email lists and bulletin boards which, in addition to being communication platforms, serve as automatic and searchable archives that members can (and should) use before posting questions. People are encouraged to use existing knowledge rather than asking other community goers for already archived information. For the above reasons, we believe that studies dealing with the determinants of knowledge contributions might benefit from taking interest in identifying who may possess the relevant knowledge in the field of the problems being discussed.

### 3.2. Hypotheses

As mentioned above, lead users have certain characteristics that may play a key role in shaping their knowledge sharing behavior in a community. The three defining characteristics of lead users are that they are early adopters, that they perceive the need for an innovation earlier than others, and that they expect great benefit from solving innovation-related problems in this area. These three characteristics are conjectured to shape the way in which these users share knowledge in the community context.

Our first hypothesis pertains to lead users as knowledge contributors to the community. As one of the three defining characteristics of lead users is that they are early adopters of new products or services, these users can be assumed to have experience of a given technology or question early on. As lead users are ahead of the general practice of the community in terms of use and development of the product in question, they will be more likely to hold the answers to the bulk of the questions that ordinary members post in the community. Given the fact that online communities work under the norm that mainly up-to date questions that are not to be found in the archive are posted, it is more likely that discussions will be of recent issues, thereby increasing the likelihood of lead users holding the answer. We argue that given the characteristics of lead users, there will be a higher likelihood of a match-up between a question from another community member and the knowledge base of lead users. Lead users will be more likely to be able to search and match (Olivera et al., 2008) what they perceive to be the answers needed and will therefore contribute more intensely. Accordingly, we hypothesize:

**H1.** Lead user characteristics are positively related to user propensity to give knowledge to the community.

Searching inspiration for a solution locally may not always yield the appropriate outcomes due to the limited scope of the local knowledge base. Our second hypothesis relates to lead users as individuals with the incentive to seek inspiration in different communities and to disseminate the derived knowledge within the focal community. This activity can be illustrated as gatekeeping (Allen, 1977), processes by which individuals bring about some of the variety necessary to create "new combinations." The existence of the "gatekeeper," an individual who links his or her organization to the world at large, was pointed to by Allen and Cohen (1969).

In dealing with novel self-defined problems in the process of solving needs ahead of the trend curve where no solutions are yet available (where lead users are active), all the knowledge inputs required by the lead users will likely not coincide with their own knowledge base. In such situations, the need to go beyond the boundaries of a present community for complementary knowledge may be quite prevalent (Postrel, 2002). Hence, local sources of inputs may often offer too little inspiration and variety to solve a lead user's problems. Here, we argue that boundary-spanning activities (here in terms of getting inspiration from other communities) enhance individuals' capacity for problem-solving. Accordingly, we expect a complementary relationship between seeking knowledge in other communities and having a high lead user characteristics

score to affect the proclivity of the user to give knowledge to the community. In sum, we posit:

**H2.** The number of communities the user engages in positively moderates the relationship between lead user characteristics and the inclination of the user to give knowledge to the community.

While paying attention to earlier literature on the relationship between motivations and knowledge contribution in communities, we find that the theory is not complete without actually testing for the coexistence between our perspective and the perspective on motivation and contribution. We would expect that these two perspectives are additive and not substitutes. Therefore, our expectation is, in line with earlier studies, that intrinsic and extrinsic types of motivation will be positively related to making contributions in the community.

## 4. Empirical analyses

### 4.1. Methodology

Our analysis is based on results from a survey in combination with data obtained from web logs generated by active users in an online community and backed by interviews with key respondents. We sent out a web-based questionnaire to the community of users located in Propellerhead Software's community, Propellerhead Community. The questionnaire was launched on May 14, 2003 and continued through to June 18, 2003. The objective of the survey was to collect data on users' personal characteristics, particularly those regarding lead users. The object studied, innovative users located in an online community, favored the use of a web-based survey method. The community goers were asked questions about: their background, community participation information, innovative work carried out, motivations for community participation, and knowledge sharing inside and outside the community. The questionnaire appeared in a pop-up window when a community participant logged into the online community (posted at the Internet location Propellerheads.se). When finished, the respondent submitted the questionnaire directly to our database.

The questionnaire had a response rate of 62.7 percent (i.e. 62.7 percent of those offered the questionnaire responded). The total number of responses was 442. The first best choice for avoiding non-response bias is to achieve a high response rate (Armstrong and Overton, 1977)—a 63 percent response rate must be considered very good in this respect. With respect to common method bias, the questionnaire was constructed in such a way that the respondents had to answer in different ways in various parts of the questionnaire. In other words, we used a mix of Likert scales, yes/no questions, estimates of time periods spent, etc., throughout the questionnaire. Moreover, we combined the questionnaire-based data with simple measures of actual behaviors in terms of posting activities in the online community (more about this in Section 4.2 below). Finally, we performed Harman's one-factor test on the items included in our regression model to examine whether common method bias may augment the relationships detected. Since we found multiple factors, and since the first factor does not account for the majority of the variance (the first factor accounts for only 25 percent of the variance), potential problems associated with common method bias are not indicated by the test (Podsakoff and Organ, 1986).

Our dependent variables are ordered response variables (see, Wooldridge, 2002: 504–508, for an exposition of models dealing with ordered responses). As the name suggests, if the dependent variable, $y$, is an ordered response, then the values we assign to each outcome are not arbitrary. In our case, the dependent variables are measured on Likert scales. For instance, our dependent

variable "knowledge give" is on a scale from one to seven, with $y = 7$ representing the highest ranking and $y = 1$ the lowest ranking. The fact that seven is a higher ranking than six conveys useful information, even though the ranking itself has ordinal meaning only. In particular, we cannot say that the difference between four and two is somehow twice as important as the difference between two and one—this is the assumption one would have to make to use ordinary least squares (OLS) estimation. When the dependent variable is an ordered response variable, the ordered probit or ordered logit models are appropriate econometric estimation techniques (in this paper we report ordered probit estimates only, but the ordered logit estimations are strikingly similar and are available upon request). The relevant parameters for obtaining response probabilities for each of the outcomes of the dependent variable can be estimated using maximum likelihood. The precise log-likelihood function is given by Wooldridge (2002: 505).

We use Stata 10.0 to obtain our ordered probit estimates. The sign of the effect of the independent variables on the probability of the lowest outcome $P(y = 1|\mathbf{x})$ and the highest outcome $P(y = 7|\mathbf{x})$ is given by the estimate of the coefficient ($\mathbf{x}$ is a vector of explanatory variables). However, the sign of the effect on the intermediate categories is not unambiguously determined by the coefficient estimate. In order to get an idea of the sign and magnitude of these effects, we also – following the convention – report marginal effects (at the mean) in Appendix Tables A.1 and A.2 for most of our ordered probit estimations.

### 4.2. Variables

#### 4.2.1. Dependent variables

The key behavioral dependent variable measures whether the user enjoys giving assistance as an expert ("knowledge give"). The variable is based on the following question on the questionnaire: "Check out the statements below and indicate on the 1 to 7 point scale if they fit your characteristics: I enjoy giving others advice as an expert." Our reason for choosing a research design involving posing the question in terms of enjoying giving others advice as an expert is motivated by the fact that asking users directly for how much knowledge they give to the community would by too subjective to be of use. Accordingly, we assume that when users like to give knowledge as an expert, they will also be inclined to actually share the knowledge they possess.

However, in order to validate our "knowledge give" variable, we both correlated our variable with a simple and objective measure of actual behavior in the online community and used the objective measure as a dependent variable in some of our regression models. We got access to all postings (31,125 postings) by members of the community over the period February 14, 2002 to March 9, 2003. We asked two persons to rate whether the given members, in 1600 postings to the website (one person rated 700 postings and another rated 900 postings): (1) posed a question, (2) gave an answer, or (3) wrote something that was neither (1) nor (2). In order to check for inter-coder reliability, we had 100 postings double-rated by the two coders. We then calculated the coefficient of reliability and found that the two coders were in agreement 88 percent of the time. In order to account for agreement due to chance, we also calculated Scott's pi and found a value of 0.87, which is typically considered to indicate very good inter-coder reliability. Based on the ratings of the postings, we calculated a variable expressing the observed propensity to post answers (percentage of answers as a proportion of total postings) in the community. Since we did not have all postings rated over the period, we were only able to identify 154 individuals who had also filled out the questionnaire with non-missing observations for all our variables. However, this variable correlates positively with the "knowledge give" vari-

able (0.1 percent level of significance and a correlation coefficient of 0.4).[2]

#### 4.2.2. Key independent variables

Our key independent variable is lead user characteristics: it is based on the lead user construct (Morrison et al., 2000) and involved three questions that identify leading edge users: (1) "I usually find out about new products and solutions earlier than others", (2) "I have benefited significantly from early adoption and use of new products" and (3) "I have tested prototype versions of new products for manufacturers to a large extent." As indicated above, each of these questions could be answered using a seven-point Likert scale. The three items were then collapsed into one single variable by means of summation. As our key independent variable – lead user – was constructed from those three items we chose to perform a standardized Cronbach's Alpha test. Since the standardized Cronbach's Alpha is 0.7, the variable has an acceptable degree of internal validity. Our decision to treat "lead user" as a non-dichotomous variable was motivated by some recent evaluations of the construct which found the non-dichotomous analog to its binary ancestor more appropriate due to its ability to deal with degrees of lead user characteristics (see, Morrison et al., 2004). Fig. 1 plots frequencies of the index of lead user characteristics. It can be seen that the distribution is bell-shaped, with a tendency for the peak to be on the right hand side of the distribution. Our moderator variable, "number of communities", is based on the following question on the questionnaire: "How many communities related to sound production and modding of virtual instruments do you participate in?" The values of the variable range from 1 to 9.

#### 4.2.3. Control variables

An important control variable is user innovation ("innovator"). The variable is discrete and is constructed as follows. To test whether users had innovated, they were asked: "Have you developed modifications, add-ons or extras to Propellerhead's products? yes/no." By asking in this manner we were able to establish innovation as a discrete variable: if users had innovated, the innovation variable was set to the value of 1 and if no innovation was reported, the value of the innovation variable was set to equal 0. We asked respondents who answered yes to having innovated to provide us with evidence of their innovation via electronic mail. This way we were able to verify the existence of 26 out of 34 innovations and thus that individuals reporting having created an innovation had, in fact, done so. Another important control variable is a "user experience" measure and is based on the following question: "I have been using Propellerheads products for: (a) Less than 1 month, (b) Less than 6 months, (c) Less than 1 year, (d) Between 1 and 4 years, or (e) More than 4 years." We transformed these categories into "continuous" numbers, with (a)–(e) becoming 0.5, 3, 9, 24 and 72 months, respectively (the results are not sensitive to this precise transformation). However, it may not only be the time period of using the product but also the intensity of use that should define "user experience". Therefore, we required that users should spend at least 5 h a week on average on "activities related to sound production or processing" in order not to get the lowest "user experience" score. Another control variable is "focus on technological product improvements". The variable is discrete and is based on the following question: "I spend most of my time on technological product improvements, making mods, yes/no." Another set of control variables is based on a question pertaining to the primary occupation of the community member: "I spend most of my time as:

---

[2] We also correlated the propensity to post answers with our lead user construct and found a positive and significant (1 percent level of significance) relationship between the two variables (correlation coefficient of 0.2).
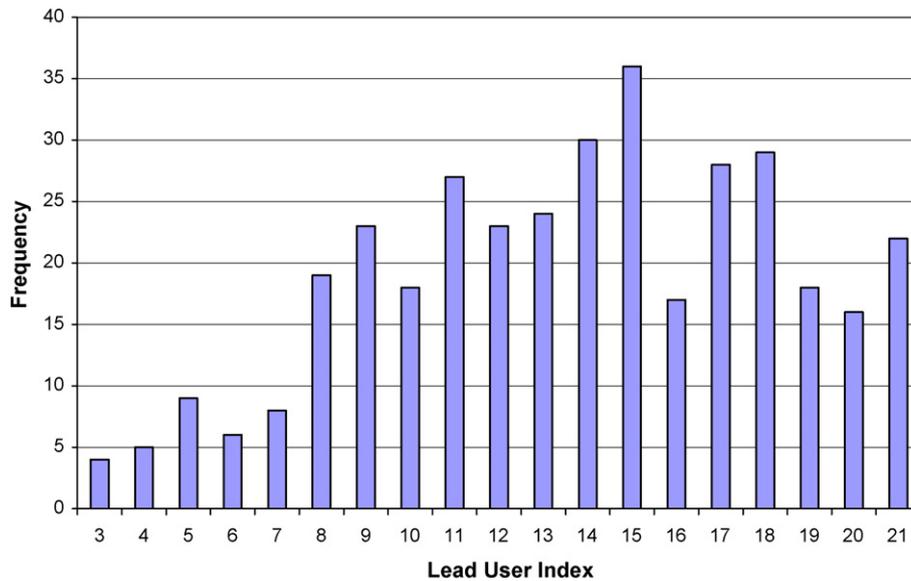
**Fig. 1.** Frequencies of index of lead user characteristics.

(a) professional programmer, (b) hobby programmer, (c) student, and (d) other". On the basis of this question, we created three binary variables ("professional programmer", "hobby programmer", "student") and used the "other" category as the benchmark in the estimations. The "peer recognition important" variable is based on the question: "Check out the statements below and indicate on the 1–7 point scale whether they fit your characteristics: Recognition from other community goers is my greatest reward", while the "do it for fun" control variable draws on the question: "What are the reasons for your participation in the community? For fun, yes/no".

### 4.2.4. Descriptive statistics

Table 1 gives descriptive statistics for the variables used in this study. It can be seen that the average lead user score is about 14, while the maximum value is 21 (corresponding to the theoretical maximum). The strongest correlation between the independent variables is between whether the users judge peer recognition to be important as a motivating factor on the one hand and enjoy giving knowledge as an expert on the other ($r = 0.26$).

### 4.3. Regression results

Table 2, Model 2 contains the estimations of the ordered probit model with all key independent and control variables included. The corresponding marginal effects at the means for each of the seven outcomes of the dependent variable are reported in Appendix Table A.1.

According to Model 2, lead user characteristics appear to relate positively to inclination of the user to give knowledge to the community (as conjectured in Hypothesis 1) since the pertinent parameter is significant at the one percent level as well. Moreover, the marginal effects are negative for KG = 1 . . . 5 and positive for KG = 6 and 7 (see Appendix Table A.1). Stated differently, lead user characteristics decrease the probability of knowledge give when the rating is below six and increase the probability of knowledge give when the rating is six or seven on our seven-point scale. Again, the estimations are well-behaved as the highest positive marginal effect is found for KG = 7. In order to test the robustness of this finding we replaced the "knowledge give" variable with the observed propensity to post answers to the community. The results of this experiment are reported in Model 3. We find that lead user characteristics are a positive and significant parameter (one percent level) in explaining the proportion of answers over total postings. Although we only had 154 observations when we performed this analysis, we find that results are similar to the results produced when using the ordered probit and the entire sample in Model 4. Moreover, when we ran the ordered probit model using knowledge give as the dependent variable while focusing on the 154 observations for which we have the postings measure, we got similar results (these results are not given for reasons of space).

**Table 1**
Descriptive statistics ($n = 364$).

| Variable | Mean | Std. dev. | Min | Max | 1. | 2. | 3. | 4. | 5. | 6. | 7. | 8. | 9. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. Enjoy giving assistance as an expert ("knowledge give") | 4.87 | 1.79 | 1.00 | 7.00 | | | | | | | | | |
| 2. Lead user characteristics | 0.10 | 4.50 | −10.54 | 7.46 | 0.25 | | | | | | | | |
| 3. Number of communities | 0.07 | 1.61 | −1.43 | 6.57 | 0.04 | 0.18 | | | | | | | |
| 4. Innovator | 0.09 | 0.28 | 0.00 | 1.00 | 0.18 | 0.15 | 0.01 | | | | | | |
| 5. Log user experience | 0.26 | 20.39 | −14.07 | 57.93 | 0.20 | 0.30 | 0.14 | 0.04 | | | | | |
| 6. Focus on technological product improvements | 0.02 | 0.13 | 0.00 | 1.00 | −0.03 | 0.11 | 0.04 | −0.04 | 0.04 | | | | |
| 7. Professional programmer | 0.16 | 0.36 | 0.00 | 1.00 | 0.02 | 0.07 | 0.01 | −0.05 | −0.07 | 0.06 | | | |
| 8. Student | 0.15 | 0.36 | 0.00 | 1.00 | 0.00 | −0.07 | −0.03 | 0.07 | 0.00 | 0.01 | −0.14 | | |
| 9. Peer recognition important | 0.00 | 1.98 | −2.53 | 3.47 | 0.26 | 0.16 | 0.00 | 0.09 | 0.06 | 0.02 | −0.01 | 0.12 | |
| 10. Do it for fun | 0.02 | 0.47 | −0.66 | 0.34 | 0.11 | 0.07 | −0.02 | 0.09 | −0.04 | 0.00 | 0.06 | −0.04 | 0.16 |

*Note*: Correlations equal to or above |0.10| are significant at $p < 0.05$. Two-tailed tests.

**Table 2**
Regressions, explaining knowledge sharing behavior in the Propellerhead online community.

| Dependent variables | Model 1 Ordered probit Enjoy giving assistance as an expert ("knowledge give") | | Model 2 Ordered probit Enjoy giving assistance as an expert ("knowledge give") | | Model 3 OLS Number of answers as a proportion of total postings | | Model 4 Ordered probit "Enjoy giving assistance as an expert ("knowledge give")" | | Model 5 OLS "Enjoy giving assistance as an expert ("knowledge give")" | |
|---|---|---|---|---|---|---|---|---|---|---|
| Independent variables | Coef. | SE | Coef. | SE | Coef. | SE | Coef. | SE | Coef. | SE |
| Lead user characteristics | 0.055*** | 0.013 | 0.038*** | 0.014 | 0.067*** | 0.028 | 0.040*** | 0.014 | 0.061*** | 0.021 |
| Lead user characteristics × number of communities | | | | | | | 0.014** | 0.008 | 0.018* | 0.012 |
| Number of communities | −0.005 | 0.035 | −0.009 | 0.035 | −0.103* | 0.067 | −0.028 | 0.037 | −0.040 | 0.059 |
| Innovator | 0.623*** | 0.209 | 0.574*** | 0.212 | −0.290 | 0.373 | 0.550*** | 0.212 | 0.817*** | 0.321 |
| Log user experience | | | 0.009*** | 0.003 | 0.003 | 0.005 | 0.009*** | 0.003 | 0.012*** | 0.005 |
| Focus on technological product improvements | | | −0.408 | 0.438 | −0.692 | 1.460 | −0.503 | 0.441 | −0.774 | 0.697 |
| Professional programmer | | | 0.065 | 0.155 | −0.042 | 0.382 | 0.064 | 0.155 | 0.094 | 0.246 |
| Student | | | −0.069 | 0.160 | −0.084 | 0.311 | −0.052 | 0.160 | −0.087 | 0.253 |
| Peer recognition important | | | 0.125*** | 0.030 | 0.110** | 0.059 | 0.128*** | 0.030 | 0.194*** | 0.046 |
| Do it for fun | | | 0.123 | 0.120 | −0.231 | 0.250 | 0.129 | 0.121 | 0.236 | 0.191 |
| Constant | | | | | 0.890*** | 0.139 | | | 4.772*** | 0.109 |
| No. of obs. | 364 | | 364 | | 154 | | 364 | | 364 | |
| Chi-square | 32*** | | 63*** | | | | 66*** | | | |
| McFadden's pseudo R² | 0.02 | | 0.05 | | | | 0.05 | | | |
| R² | | | | | 0.09 | | | | 0.16 | |

*Note*: One-tailed test applied.
* p < .10.
** p < .05.
*** p < .01.

With respect to our moderating hypothesis, Hypothesis 2 ("The number of communities the user engages in positively moderates the relationship between lead user characteristics and the inclination of the user to give knowledge to the community"), we find support for this hypothesis given the positive and significant parameter (five percent level) for the interacted variable, lead user characteristics × number of communities, found in Table 2, Model 4, and given that associated marginal effects at the mean are negative for KG = 1 . . . 5 and positive for KG = 6 and 7 (see Appendix Table A.2). Put differently, lead user characteristics × number of communities decreases the probability of knowledge give when the rating is below six and increases the probability of knowledge give when the rating is six or seven on our seven-point scale. The highest positive marginal effect is found for KG = 7. Given that interaction effects in the ordered probit model are notoriously difficult to interpret because the estimated probabilities depend on the values of the other covariates (Long and Freese, 2006), we also estimated the model using OLS, although OLS is an inefficient estimator in this case. The results of this estimation can be found in Table 2, Model 5. Although the interaction effect is not as significant as that found when using the more appropriate ordered probit estimator – it is significant at the 10 percent level only, as opposed to the 5 percent level for the ordered probit estimator – the OLS results are nevertheless consistent with the ordered probit results.

Through our post hoc validation exercise we found a general agreement between the statements from interviews with two lead user respondents and our main results. They both found knowledge give an important part of day-to-day practice in the community. They also stated that it would be natural to share the information gained from interaction in other communities related to computer-controlled music instruments – and from communities interested in topics beyond these specific technologies – with individuals in the Propellerhead community when necessary and when asked.

## 5. Conclusion and discussion

In this paper, we have analyzed knowledge sharing behaviors in an online community and in this context put the role of lead users under scrutiny. We proposed a model portraying lead users as problem solvers with a high likelihood of being able to provide knowledge to others in the community and who would actively do so, given the low cost of providing knowledge already at hand. We also argued that lead users would be likely to perform a boundary-spanning and gate keeping role, exposing them to novel sources of valuable knowledge to their activities in the focal community. The model found support in our empirical results since we have shown that users with a high degree of lead user characteristics tend to enjoy revealing their knowledge to other users. Moreover – and still in agreement with our model – we found that users with many lead user characteristics, who also conduct boundary-spanning across several online communities, are more prone to like sharing knowledge in the online community.

Our results also indicate that users in general do not tend to span the boundaries of the community as lead users do. This may also be a sign of lesser demand for new knowledge and be a sign of engagement in less demanding innovation processes. Throughout the models, the motivation peer recognition is positive and significantly related to knowledge give; thus underscoring the importance of extrinsic motivation in knowledge sharing that prior studies have highlighted (Wasko and Faraj, 2005). However, the "fun" motivator was positive, but not significant. Thus, in sum, there seem to be some coexistence of at least extrinsic motivation for

making contributions with the perspective that we have outlined in this paper.

We have added to the growing body of literature on the phenomenon of online communities and the relationship between key community members and their participation in knowledge sharing processes. Knowledge sharing is an important aspect of innovative communities of practice and a necessary condition for the ability of communities to become innovative in the first place. Prior research on online communities of practice has put under scrutiny a range of different motive for users to contribute to such communities. We have added to that view by proposing – and finding support for – the idea that in a community characterized by posting and by problems presented in the community being ahead of the archive, we can explain knowledge contributions as originating from the sources which are likely to hold that knowledge. We have paid attention to particular information technologies of online communities of practice that enable posting and ensure that problems are ahead of the archive. Thus, our findings may only be generalized to settings in which these modes of communications are present.

Theory on lead users suggests that such users are related to innovation. Our study suggests that lead users are also a potentially interesting unit of analysis for explaining how new knowledge is integrated into communities and how it is diffused locally in the community. By focusing on lead users' role in knowledge sharing, we also added to the literature on online communities of practice. Our study has shown that in making predictions about the sources of knowledge provision, it may be fruitful to focus on those particular individuals who can be expected to hold knowledge relevant for sharing.

We suggest that the fact that lead users span the boundaries of the community combined with a high level of sharing inside the community may well explain why online communities of practice may in many case also be important innovative communities in their area of practice. If, on the contrary, the most leading edge individuals in the community were always preoccupied with activities at the center of their fields (as in Kuhnian "normal science" 1970) they would fail to be exposed to ideas other than those already accepted and known in the field/community. Our finding related to boundary-spanning activity indicates that lead users may be likely sources of adjustments, recombinations or extensions of existing principles within the field and thus provide novel inputs to the community. Lead users tend not only to find and absorb knowledge outside the community; importantly, our findings also indicate that they share it inside the community. Thus, lead users may be described in a social setting as not only the lead adopters that sometimes innovate, but also as gatekeepers, spanning the boundaries of the community and as linking their community to the world at large.

The findings raise questions about how managers should confront online communities of practice. In the context of a firm-hosted community of practice, in which firms depend on input from innovative members of their community, managers and community moderators must realize that trying to identify and retain lead users is centrally important. Accordingly, managers and community moderators need to think about how they can motivate users not only to make contributions to the community, but also to identify key players and to create incentives for users to stay and keep on contributing.

## Appendix A. Marginal effects

.

**Table A.1**
Marginal effects at the mean for Model 2.

| Variable | KG = 1 | KG = 2 | KG = 3 | KG = 4 | KG = 5 | KG = 6 | KG = 7 |
|---|---|---|---|---|---|---|---|
| Lead user characteristics | −0.004[***] | −0.002[***] | −0.004[***] | −0.005[***] | 0.000 | 0.003[***] | 0.012[***] |
| Number of communities | 0.001 | 0.001 | 0.001 | 0.001 | 0.000 | −0.001 | −0.003 |
| Innovator | −0.041[***] | −0.027[***] | −0.049[***] | −0.085[***] | −0.024[*] | 0.026[***] | 0.200[***] |
| Log user experience | −0.001[***] | −0.001[***] | −0.001[***] | −0.001[***] | 0.000 | 0.001[***] | 0.003[***] |
| Focus on technological product improvements | 0.059 | 0.027 | 0.039 | 0.037[**] | −0.014 | −0.042 | −0.105 |
| Professional programmer | −0.007 | −0.004 | −0.006 | −0.008 | 0.000 | 0.005 | 0.020 |
| Student | 0.008 | 0.004 | 0.007 | 0.009 | 0.000 | −0.006 | −0.021 |
| Peer recognition important | −0.013[***] | −0.007[***] | −0.012[***] | −0.016[***] | 0.000 | 0.011[***] | 0.038[***] |
| Do it for fun | −0.013 | −0.007 | −0.012 | −0.016 | 0.000 | 0.010 | 0.037 |

*Note*: One-tailed test applied.
[*] $p < .10$.
[**] $p < .05$.
[***] $p < .01$.

**Table A.2**
Marginal effects at the mean for Model 4.

| Variable | KG = 1 | KG = 2 | KG = 3 | KG = 4 | KG = 5 | KG = 6 | KG = 7 |
|---|---|---|---|---|---|---|---|
| Lead user characteristics | −0.004[***] | −0.002[***] | −0.004[***] | −0.005[***] | 0.000 | 0.003[***] | 0.012[***] |
| Lead user characteristics × Number of communities | −0.001[**] | −0.001[*] | −0.001[**] | −0.002[**] | 0.000 | 0.001[**] | 0.004[**] |
| Number of communities | 0.003 | 0.002 | 0.003 | 0.004 | 0.000 | −0.002 | −0.009 |
| Innovator | −0.040[***] | −0.026[***] | −0.048[***] | −0.081[***] | −0.022[*] | 0.026[***] | 0.191[***] |
| Log user experience | −0.001[***] | −0.001[***] | −0.001[***] | −0.001[***] | 0.000 | 0.001[***] | 0.003[***] |
| Focus on technological product improvements | 0.076 | 0.034 | 0.047 | 0.040[***] | −0.021 | −0.054 | −0.124[*] |
| Professional programmer | −0.006 | −0.004 | −0.006 | −0.008 | 0.000 | 0.005 | 0.020 |
| Student | 0.006 | 0.003 | 0.005 | 0.006 | 0.000 | −0.004 | −0.015 |
| Peer recognition important | −0.013[***] | −0.008[***] | −0.012[***] | −0.016[***] | 0.000 | 0.011[***] | 0.039[***] |
| Do it for fun | −0.013 | −0.008 | −0.012 | −0.016 | 0.000 | 0.011 | 0.039 |

*Note*: One-tailed test applied.
[*] $p < .10$.
[**] $p < .05$.
[***] $p < .01$.

# References

Allen, T.J., 1977. Managing the Flow of Technology. MIT Press, Cambridge, MA.

Allen, T.J., Cohen, S.I., 1969. Information flow in research and development laboratories. Administrative Science Quarterly 14, 12–19.

Armstrong, A., Hagel, J., 1996. The real value of on-line communities. Harvard Business Review 74, 134–141.

Armstrong, J.S., Overton, T.S., 1977. Estimating non-response bias in mail surveys. Journal of Marketing Research 14, 396–402.

Benkler, Y., 2004. Intellectual property: commons-based strategies and the problems of patents. Science 305, 1110–1111.

Borgatti, S.P., Cross, R., 2003. A relational view of information seeking and learning in social networks management science 49, 432–445.

Brown, J.S., Duguid, P., 2000. The Social Life of Information. Harvard Business School Press, Boston, MA.

Constant, D., Kiesler, S., Sproull, L., 1996. The kindness of strangers: the usefulness of electronic weak ties for technical advice. Organization Science 7, 119–135.

Edmondson, A., 1999. Psychological safety and learning behavior in work teams. Administrative Science Quarterly 44, 350–383.

Finholt, T., Sproull, L., 1990. Electronic groups at work. Organization Science 1, 41–64.

Franke, N., Shah, S., 2003. How communities support innovative activities: an exploration of assistance and sharing among end-users. Research Policy 32, 157–178.

Gallivan, M.J., 2000. Examining workgroup influence on technology usage: a community of practice perspective. In: Paper Presented at the Proceedings of the ACM SIGCPR Conference on Computer personnel Research.

Goodman, P.S., Darr, E.D., 1998. Computer-aided systems and communities: mechanisms for organizational learning in distributed environments. MIS Quarterly 22, 417–440.

Haefliger, S., von Krogh, G., Spaeth, S., 2008. Code re-use in open source software. Management Science 54, 180–193.

Hansen, M.T., 1999. The search-transfer problem: the role of weak ties in sharing knowledge across organizational subunits. Administrative Science Quarterly 44, 82–111.

Hertel, G., Niedner, S., Herrmann, S., 2003. Motivation of software developers in Open Source projects: an Internet-based survey of contributors to the Linux kernel. Research Policy 32, 1159–1177.

Jeppesen, L.B., Frederiksen, L., 2006. Why do users contribute to firm-hosted user communities? The case of computer-controlled music instruments. Organization Science 17, 45–63.

Kogut, B., Zander, U., 1992. Knowledge of the firm, combinative capabilities, and the replication of technology. Organization Science 3, 383–397.

Kollock, P., 1999. The economies of online cooperation: gifts and public goods in cyberspace. In: Smith, M., Kollock, P. (Eds.), Communities in Cyberspace. Routledge, London.

Kuhn, T.S., 1970. The Structure of Scientific Revolutions. University of Chicago Press, Chicago/London.

Lakhani, K., Wolf, R., 2005. Why Hackers do what they do: understanding motivation and effort in free/open source software projects. In: Feller, J., Fitzgerald, B., Hissam, S., Lakhani, K. (Eds.), Perspectives in Free and Open Source Software. MIT Press, Cambridge, MA.

Lakhani, K.R., von Hippel, E., 2003. How open source software works: "free" user-to-user assistance. Research Policy 32, 923–943.

Lave, J., Wenger, E., 1991. Situated Learning: Legitimate Peripheral Participation. Cambridge University Press, Cambridge, England.

Lerner, J., Tirole, J., 2002. Some simple economics of open source. Journal of Industrial Economics 50, 197–234.

Long, J.S., Freese, J., 2006. Regression models for categorical dependent variables using Stata. Stata Press, College Station, TX.

Ma, M., Agarwal, R., 2007. Through a glass darkly: information technology design, identity verification, and knowledge contribution in online communities. Information Systems Research 18, 42–67.

Majchrzak, A., Cooper, L.P., Neece, O.E., 2004. Knowledge reuse for innovation. Management Science 50, 174–188.

Markus, L.M., 2001. Towards a theory of knowledge reuse: types of reuse situations and factors in reuse success. Journal of Management of Information Systems 18, 57–93.

Moon, J.Y., Sproull, L., 2000. The essence of distributed work: the case of the Linux kernel. In: First Monday 5., http://ftp.firstmonday.dk/www/issues/issue5_11/moon/.

Morrison, P.D., Roberts, J.H., Midgley, D.F., 2004. The nature of lead users and measurement of leading edge status. Research Policy 33, 351–362.

Morrison, P.D., Roberts, J.H., von Hippel, E., 2000. Determinants of user innovation and innovation sharing in a local market. Management Science 46, 1513–1527.

Muniz, A.M., O'Guinn, T.C., 2001. Brand community. Journal of Consumer Research 27, 412–432.

Nonaka, I., 1994. A dynamic theory of organizational knowledge creation. Organization Science 5, 14–37.

Olivera, F., Goodman, P.S., Tan, S.S.-L., 2008. Contribution behaviors in distributed environments. MIS Quarterly 32, 23–42.

Orr, J.E., 1996. Talking About Machines: An Ethnography of a Modern Job. ILR Press, Itchaca, NY.

Podsakoff, P.M., Organ, D.W., 1986. Self-reports in organizational research: problems and prospects. Journal of Management 12, 531–544.

Postrel, S., 2002. Islands of shared knowledge specialization and mutual understanding in problem-solving teams. Organization Science 13, 303–320.

Raymond, E.S., 1999. The Cathedral and the Bazaar: Musings on Linux and Open Source by an Accidental Revolutionary. O'Reilly, Cambridge, MA.

Tsai, W., Ghoshal, S., 1998. Social capital and value creation: the role of intrafirm networks. Academy of Management Journal 41, 462–476.

von Hippel, E., 1986. Lead users: a source of novel product concepts. Management Science 32, 791–805.

von Hippel, E., von Krogh, G., 2003. Open source software and the "private-collective" innovation model: issues for organization science. Organization Science 14, 209–223.

von Krogh, G., 1998. Care in knowledge creation. California Management Review 40, 133–153.

von Krogh, G., 2002. The communal resource and information systems. Journal of Strategic Information Systems 11, 85–107.

Wasko, M.M., Faraj, S., 2000. It is what one does: why people participate and help others in electronic communities of practice. Journal of Strategic Information Systems 9, 155–173.

Wasko, M.M., Faraj, S., 2005. Why should I share? Examining social capital and knowledge contribution in electronic communities. MIS Quarterly 29, 35–57.

Wooldridge, J.M., 2002. Econometric Analysis of Cross Section and Panel Data. The MIT Press, Cambridge, MA.