Profit seeking identity and productivity in open source communities

By Mekki MacAulay

Abstract

The issue of profit seeking—an identification as someone who seeks financial compensation directly or indirectly for work efforts—has been a longstanding point of contention in open source communities. Profit seeking has taken on prototypical form such that it has become a distinct group that individual participants in open source communities use for identity self-categorization. Open source communities differentiate themselves in part by their profit seeking identity, represented by the perceptions of organization members of the degree to which the community allows, encourages, and supports individual profit seekers. The present research investigates the impact of both individual- and organizational-level profit seeking identity on the productivity of open source communities. Hypotheses are formulated that suggest that profit seeking identities at both the individual and organizational level are positively correlated with individual and organizational productivity respectively. Further, it is hypothesized that organizational profit seeking identity has a cross-level positive effect on individual productivity, controlling for the individual profit seeking identity. Finally, it is hypothesized that organizational profit seeking identity moderates the effect of individual profit seeking identity on individual productivity and that this moderation is moderated by the level of agreement about an organization's profit seeking identity. A research design is proposed to test these hypotheses using survey and archival data to be collected from active open source communities. The potential implications for research and practice are discussed.

Keywords: open source, identity, entrepreneurship, social identity, organizational identity, meso organizational theory, productivity, peer production, FLOS

"So he's a guy that's trying to spin off a business from what's going on inside the community. And he's still part of the community in the sense that some of the stuff that he's doing he throws back over the fence to the Persephone project. But he's very much looking to make a business. —Malcolm Reynolds

Introduction

The emergence of open source[1] communities over the past three decades has drawn the interest of scholars as these organizations appear to not fit within the conventional models of organized production. Nearly every Fortune 500 company depends on open source software to run its business and the net economic impact of open source production worldwide is in the hundreds of billions of dollars range (MERIT, 2006). Research has investigated how open source production works as a distinct mode of organizing with loose oversight, self-selection, and shared outcomes (Demil & Lecocq, 2006). Studies have also investigated the seemingly altruistic intentions of individual-level participants who volunteer their time to develop a collective good. These motivations include fostering a sense of belonging to a community (Bonaccorsi & Rossi, 2006), enabling reciprocal support for other activities (Henkel, 2003), and using participation as a venue for showcasing talent and developing one's reputation in a community (Lerner & Tirole, 2001). At the organizational level, research has investigated open source communities and their properties, success factors, pitfalls, and appropriate research measures (Crowston, Howison, & Annabi, 2006). The success of the communities that surround major open source projects is of economic importance for large corporations as the continued

---

[1] Some communities have strong feelings about the choice of words used to describe them. Some insist that "free", in the sense of "liberty" as opposed to "at no cost", is the only apt description. Others prefer the use of "open" to avoid ambiguity. These discussions, while important to the communities in question, are beyond the scope of this paper. The term open source is used in this paper for simplicity. It should be read as a general term that encompasses all forms of free/libre/open source common goods and methods and not interpreted as referring to a specific model of production, development method, or license.

development and availability of the products the projects produce has a direct impact on corporate operations. Successful open source communities sustain or grow contribution levels, number of participants, and frequency of code base modification and adaptation over time (von Krogh, Spaeth, & Lakhani, 2003). Despite the breadth of both individual-level and organizational-level research into open source communities, to my knowledge there have been no multilevel studies of open source communities that consider the factors that account for community success across the individual and organizational levels. This paper aims to fill that gap in the literature by examining the cross-level effects of identity on productivity in open source communities.

Identity

I define individual identity as the self-categorization of an individual as a member of a defined group of persons. The group can be described by its distinctiveness from other groups, its prestige, the salience of its characteristics in a specific context, and the factors that are associated with becoming a member of that group (Ashforth & Mael, 1989). Individuals "cognitively represent the defining and stereotypical attributes of groups in the form of prototypes, [which] capture the context-dependent features of group membership" (Hogg & Terry, 2000: 123). As such, it is the cognitive representation of the prototype of a group that an individual uses when engaging in the comparison process that leads to self-categorization as either a member of the group, or not (Albert & Whetton, 1985). When a newcomer joins an organization, she engages in symbolic interactions with members of the organization, which prompts a "genesis of meaning and identity", a start of the comparison, evaluation of group prototypes, and a self-categorization process (Reichers, 1987). Implicit in this process is the

notion that by self-categorizing as a member of a particular group, the individual is stating that she views herself as having a sufficient number of the prototypical characteristics of that group to feel like the description of membership in the group is accurate.

Identity has often been examined as membership in a specific group such as a profession (Pratt, Rockmann, & Kaufmann, 2006), a race (Ibarra, 1995), a gender (Ely, 1994), or religious group (Creed, DeJordy, & Lok, 2010). Individuals can maintain identities as members of different groups. These multiple identities sometimes come into conflict which can be a motivator for action on the part of the individual who is compelled to resolve the identity conflict (Gutierrez, Howard-Grenville, & Scully, 2010). As a result, the identity of individuals is likely to have a performance impact in an organizational context, making it relevant to management research and theory development. An important facet of identity is that individuals attach value and significance to particular identities in specific contexts (Ely, 1994). In a given organizational context, a particular identity might be emphasized due to the values of the organization that support or threaten one identity, while other identities might be contextually irrelevant. As such, when considering the identity formation process of individuals, cross-level effects in the organizational context need to be examined.

I define organizational identity as the central, relatively enduring, and distinctive properties of an organization as perceived by its members (Albert & Whetton, 1985). While the classical definition of organizational identity would see it as largely static, research has shown that organizations adapt their identity over time and that it can be "relatively fluid" (Gioia, Schultz, & Corley, 2000). I focus on identity as relatively enduring in that I differentiate it from

more rapidly shifting constructs such as product or business lines and day to day activities, processes, or structural organization. All of these constructs can affect the shift of identity over time, but the process of organizational identity change is tied to individual sense making and adaptation, both of which take time (Gioia, Schultz, & Corley, 2000). Like individuals, organizations can have multiple identities (Foreman & Whetten, 2002). Conflicting organizational identities can have an effect on the individuals in the organization, leading to organizational members attempting to reconcile their perception of the organizational identity and that which they believe the organizational identity should be by taking action.

Organizational identity can be a factor in firm competitive advantage (Fiol, 2001) as the match between individual and organizational identity can bind people together to stimulate action or cause blind spots that hinder the emergence of new organizational possibilities. Organizational identity influences the construction and enactment of roles within the organization, directing resources, effort, and influencing structure (Golden-Biddle & Rao, 1997). It has a reciprocal influence on the dynamic interactions among organizational stakeholders (Scott & Lane, 2000). Given the bidirectional influence, an understanding of the impact of organizational identity on an organization's performance necessarily requires cross-level investigation to fully capture the effect.

Organizational-level and individual-level identity are largely isomorphic because there is internal coherence of goals between both organizational- and individual-level prototypes and their ability to motivate action to reconcile any difference between conflicting identities. Further, the core, distinctive, and relatively enduring aspects of nested identities tend to be

isomorphic as the cross-level effects shape the identities in both directions to try to bring them into alignment with one another. Impediments to isomorphism of identity across level can lead to differentiation or the formation of new identities at one level to resolve the conflict (Ashforth, Rogers, & Corley, 2011).

Profit seeking identity

In this study I examine profit seeking identity (PSI). I define PSI at the individual level as the self-categorization as someone who seeks to get financial compensation directly or indirectly for work efforts. It is a form of invisible identity (Clair, Beatty, & Maclean, 2005) in that it is an inward-facing identity that requires self-awareness of categorization. When someone else is evaluating a given individual, unlike gender or race, it cannot be readily determined whether or not the focal individual is a profit seeker without inferring this identity from patterns of actions and interactions with other individuals.

This focus emerged from preliminary research that I approached with a grounded methodology. As part of a larger research projects, open-ended semi-structured interviews were conducted with participants in open source communities. The informants described the tensions between participants in the community that they felt were there for purely altruistic or ideological reasons and those participants who were there because they wanted to make money. At first, it seemed that the notion was perhaps a community norm (Peteraf & Shanley, 1997) or facet of organizational culture (Golden, 1992). However, as the interviews progressed, it became clear that the informants had formed both a prototype of the "profit seeker" and used this prototype to compare and categorize themselves. In short, "profit seeker" was an

identifiable group in open source communities that individual members used to define their

identity. As one informant described it:

> *"I don't share the motivations of my friends at Persephone[2]. I am the guy who is taking their product and rebadging it so that it can be used by corporations. [I'm not a member of the community] because I have profit motive."* –Simon Tam[2]

Clearly, Simon Tam identified as a member of a group that I call "profit seekers", which

he felt somehow didn't fit in with the rest of the Persephone open source community. PSI, like

other forms of identity described in the literature, influenced his behaviour in that he engaged

in activities that were congruent with the identity (Ashforth & Mael, 1989) such as focusing on

developing areas of the project that met the specific needs of his target customer base, while

the rest of the community did not consider the needs of those customers worthy of

development effort. Simon Tam also expressed several stereotypical perceptions of self

(Ashforth & Mael, 1989) that were associated with the identity such as being interested in

getting "all the money" and joining the community for reasons beyond its mission and value

statements. He also described profit seekers as distinct and differentiated from the group of

"pure" community members in a way that separated prestige levels across the identities

(Ashforth & Mael, 1989), with the non-profit seekers in the community considered to be of

higher prestige and more important to the community than the profit seekers. The result is a

form of schizo-identification (Humphreys & Brown, 2002), where Simon Tam participates and

contributes to the community while maintaining a contrasting identity. Another informant in a

different open source community saw it different and explained:

---

[2] Pseudonyms have been used for the open source communities and informants to preserve informant anonymity

"On a personal level, I've consulted for years building solutions based on open source software. [...] I benefit because I have a better profit margin than if I built it on a proprietary base. It's an odd but very workable sort of symbiotic relationship. [...] The open source project moves ahead because we're contributing back in some way [...] and the customer wins because the solution might be up to 60% less expensive." –Jayne Cobb

Both individuals described a PSI that was related, at least in part, to their involvement in the open source communities. On the one hand, Simon Tam is uncomfortable with his PSI because he perceives it as not matching his idealized prototype of those who are "insiders" of the organization. In Tam's case, there is an identity gap (Foreman & Whetten, 2002) between his PSI and that of the organization. On the other hand, Jayne Cobb sees no problem with his PSI as he views it as symbiotic with the goals of the community. In Cobb's case, there is identity confirmation (Milton & Westphal, 2005) between the individual identity and the organizational identity. It is clear that the individual PSI and the degree to which the individual feels it is acceptable in the community, in both contexts, is inherently referent to the organizational PSI, suggesting a cross-level effect.

I define organizational-level PSI as the degree to which organization members perceive that the organization allows, encourages, and supports individual profit seekers in the community. It is a shared belief about a characteristic of the organization that shapes members' interpretations of issues, roles, actions, and feelings about the course of the organization (Golden-Biddle & Rao, 1997). Profit seeking as an organizational identity is central in that it was a foundational issue for the open source movement in the 1980s (Stallman, 1999)[3]. It is relatively enduring in that it persists as an issue that concerns participants in

---

[3] A complete treatment of the history of the emergence of open source communities is beyond the scope of this paper. Suffice it to say that one of the central issues in the GNU Manifesto (Stallman, 1999), one of the

current open source communities and affects their perceptions and actions despite continuous debate over the years that has shaped the specific means of implementing the community norms associated with the identity. It is distinctive as it is used as one of the major criteria to separate open source communities from other organizational production models, especially in the minds of participants (Albert & Whetten, 1985). The PSI of open source communities has gone through a symbolization process that has shaped how open source communities have evolved, becoming a representation for some internal characteristics of specific communities as contrasted to others (Brown & Starkey, 2000). It is differentiated from individual-level PSI in that it is a set of mutual understandings amongst members of the community that defines, in part, the norms of cooperation in the community by highlighting a potential threat to the legitimacy of the organization (Peteraf & Shanley, 1997). Individual members of the community, such as informant Simon Tam, might view an organization's PSI as distinct from his own, but might also recognize the mutual understanding that he has with other members of the community about what the habitus of the organization is or should be (Alvesson, 1994).

This conceptualization of organizational PSI in open source communities fits well with the current literature on multiple-identity organizations. It is not simply a component of the organization. Rather, it allows the organization to simultaneously, or alternatively, depending on context, be a different type of organization, such as profit seeking and socially oriented. In some communities, the overlap between the multiple identities might lead to organizational identity ambiguity, where members have more difficulty interpreting or making sense of their

---

foundational documents of the open source movement, was to distinguish the production of software for profit, which was what the outgroup, Microsoft, was doing, and the production of software as a social good to benefit all people. Later, in the late 1990s, the Open Source Initiative split off from the Free Software Foundation in part because it felt that for-profit businesses had a role to play in open source software development communities.

perceptions of the organization (Corley & Gioia, 2004). In other communities, the multiple identities can lead to a sense of collective identity in the organization, where the focus is on enabling shared purpose and similar outputs (Wry, Lounsbury, & Glynn, 2011). A prominent example is the Eclipse community (Eclipse Foundation, 2012), which was founded and is still supported primarily by funding from several large technology companies, including IBM, Red Hat, Ericsson, HP, Intel, and SAP. The notion that the outcome of joint production, the Eclipse development environment, will be used by profit seekers is taken for granted in this community, and is part of the shared purpose of organizational participation.

One of the major controversies regarding organizational PSI is that some members believe that, in order to be profit seeking, the organization must compromise its principles and put profit before all else. There is concern that if organizations focus too much on making money that they will lower their commitment to social goals or ideological positions that will harm the open source movement as a whole. In some cases, the debate takes on many characteristics of idealization of the principles of open source communities, leading to the creation of an impossible standard for community members to adhere to (Brown & Starkey, 2000). In practice, many open source communities have managed to find a middle-ground for these issues, but extremes on both ends remain. It is clear that organizational PSI is a salient distinguishing factor between open source communities that may very well impact productivity.

Productivity-impacting effects are likely to bridge across organizational and individual levels. When there is conflict between organizational and individual identities, the interactions between individual and organization are affected by the identity comparison gap (Foreman &

Whetten, 2002) that exists between the individual's PSI and the organization's PSI. When there is agreement between their identities, the interactions between individual and organization are affected by the confirmation of individual and organizational identity (Milton & Westphal, 2005). In the former case, the outcome is conflict; in the latter case, the outcome is an increase in commitment (Golden-Biddle & Rao, 1997). The outcomes of both cases affect productivity.

Hypotheses

One of the most important measures of the health and success of an open source community is the productivity of its members. Communities that have productive members tend to be more active, more relevant to their stakeholders, more likely to develop projects that will have an economic impact on organizations, and will be more likely to survive and endure as active projects (Crowston, Howison, & Annabi, 2006). Open source projects also tend to gather numerous metrics about productivity, usually represented by levels of different types of contribution, making data gathering and analysis more likely to be fruitful and relevant to the performance measures that the organizations already track. Individual productivity in an open source project is defined as the measurable contributions made by the individual to the project relative to the time and effort the individual contributed. The focus is on an individual's primary open source project as participants are often involved in more than one open source project. I select the construct of individual productivity as the individual-level dependent variable for this research as it has strategic implications for both open source communities as a whole and for a broader group of stakeholders who may be interested in the effects of contribution to projects when considering how to best address their needs. Increasing the level

of individual productivity is likely to have a positive effect on the success of an open source community.

I define organizational productivity of an open source community as the total contributions made to the community's primary open source project as measured by the metrics tracked for that project relative to the number of participants in the community. I select the construct of organizational productivity as the organizational-level dependent variable for this research as there are a number of measures of contribution that are not simply aggregations of individual level contributions, even though, by definition, organizational productivity must be the result of individual productivity. The differences in the tracked measures at the organizational level may capture relevant productivity measures that are not solely represented in the individual-level measures. There may also be reasons to believe that organizations might see different types and levels of contribution when measured as a whole than what is represented at the individual level. This possibility is especially true when considering the more difficult to quantify aspects of organizational contribution such as those due to reputation effects and intra-organizational network effects. However, for reasons of manageability, in the present study, I focus on the organizational-level contribution measures that are tracked and quantified by open source projects.

The independent variables for the study are individual- and organizational-level PSI, as previously defined. It is the effect of the degree to which an individual or organization is categorized as profit seeking on measures of contribution to the community that is of primary focus in this research.

Figure 1 summarizes the hypothesized relations between the dependent and
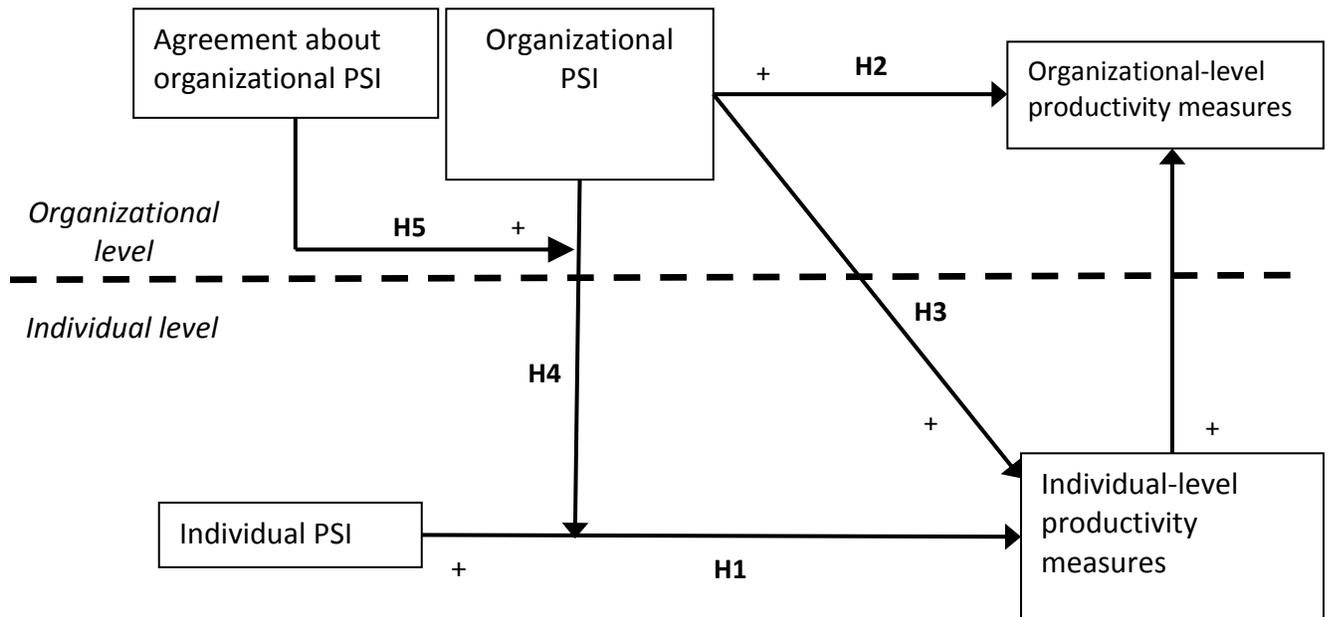independent variables at both the organizational- and individual-levels of analysis:

Figure 1: Hypothesized relationships between constructs

Individuals with a strong PSI are more likely to have entrepreneurial tendencies as
entrepreneurship combines the drive for success with an individual means of accomplishing it
in a manner that controls one's own ability to attain profitability (Shepherd & Haynie, 2009).
They are more likely to have a complementary business that can leverage the outcomes of their
productivity in the open source project (von Hippel & von Krogh, 2003). As a result, their
productivity in the open source project is part of a broader effort to make money and is less
likely to feel like wasted effort by the individual. The individual will be more productive than if
he were simply participating out of altruistic reasons, or as a casual endeavor to generate
reputation, or to foster a sense of belonging in the community (Bonaccorsi & Rossi, 2006;
Lerner & Tirole, 2002). The participants with high PSI will see their success as driven by their

own productivity which will be reflected in greater individual contributions than participants with low PSI:

Hypothesis 1: All else being equal, individual-level PSI is positively correlated with individual-level productivity

Open source communities that have a high PSI are more likely to support and encourage profit seeking activities within the community. As a result, they will attract individuals who will have profit seeking identities themselves as those individuals will select communities that provide support mechanisms and environments that match their entrepreneurial needs. Those individuals are likely to be more productive in the community than those individuals without a profit seeking identity, as per hypothesis 1. As a result, by attracting more such individuals and by facilitating their productivity, the profit seeking open source community will likely see higher levels of productivity at the organizational level:

Hypothesis 2: All else being equal, organizational-level PSI is positively correlated with organizational-level productivity

If individual PSI were the sole driver of productivity in communities, then the organizational identity's impact on productivity would be negligible. However, it stands to reason that the impact of an organization's PSI on the processes, structure, and norms of the community (Chreim, Williams, & Hinings, 2007) would result in an effect on productivity that is independent of the activities of members with specific identities. For example, an organization with high PSI may structure its productivity support and enhancement processes in such a way as to facilitate contributions from users as it knows that the net benefit to other members of the organization is significant. Non-profit seeking individuals may still be able to be productive in the organization by virtue of the impact of the organizational identity on means of producing.

The existence of such a cross-level effect of organizational identity on individual productivity follows from the notion that there is a reciprocal influence of organizational identity on individual identity formation. For individuals with initially low PSI to become individuals with high PSI, the organizational PSI must have a cross-level effect on the individuals that changes their self-categorization, and, in turn, affects and potentially stimulates their actions (Ashforth, Rogers, & Corley, 2011). The organizational dynamics between the individual and the organization over time enable new patterns of action through interpretation, legitimation, and access to material resources (Chreim, Williams, & Hinings, 2007). The impact on individual actions and interactions is likely to aid the individual in producing more effectively, separately from the impact of her own PSI. The result is effectively a meso-mediational effect on organizational productivity (Mathieu & Taylor, 2007), as organizational productivity is an aggregation of individual productivity by definition, so the influence of organizational identity on organizational productivity must mediate through individual productivity levels:

Hypothesis 3: Organizational-level PSI is positively correlated with individual-level productivity, controlling for individual-level PSI

The cross-level organizational identity effect on individual productivity is also likely to interact with individual PSI because of the degree of correspondence of individual and organizational identities affect actions. Individuals compare their identities to the identities of the organization. When the organization's identity is consistent with the individual's identity, there is identity confirmation, which builds commitment between the individual and the organization, and between the individual and other participants in the organization's social network, increasing overall individual productivity (Milton & Westphal, 2005). When the

identities do not match, there is an identity gap a conflict of commitment emerges (Golden-Biddle & Rao, 1997) and individuals will engage in congruence-enhancing responses. If the gap is small, individuals might try to bridge the gap by changing themselves or the organization. However, if the gap is too large, individuals might revaluate their relationship with the organization, distancing themselves and decreasing productivity in the organization (Foreman & Whetten, 2002). As such, the moderating effect of organizational-level PSI is positive when individual PSI is low and negative when individual PSI is high. Figure 2 describes the impact of the match of organizational and individual profit seeking identities on individual productivity the organization.
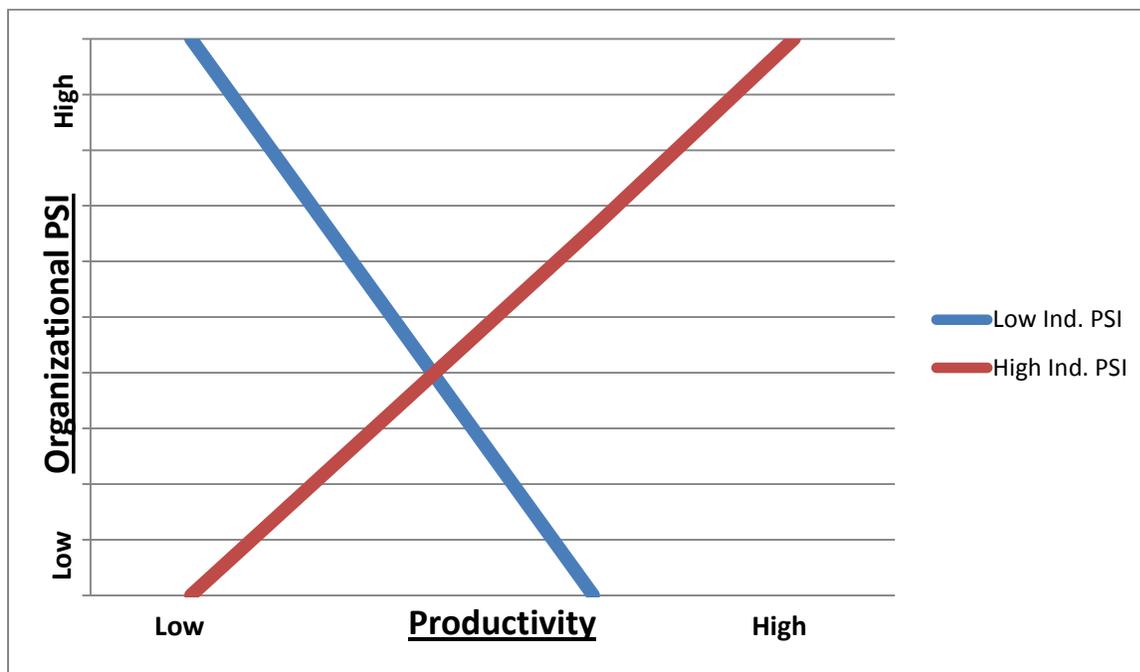

Figure 2: Effect of gap between organizational and individual PSI on individual productivity

As individual PSI is theorized to have a direct, within-level effect on individual-level productivity, all else being equal (H1), and the organizational-level PSI effect must exert itself across levels, it is theorized that high individual-level PSI will be more important for productivity

than organizational-level PSI. This difference in effect is reflected in the slope difference between low and high individual-level PSI in figure 2. The interaction effect of organization-level and individual-level PSI can be conceptualized as a skewed quadratic moderation effect of organizational-level PSI on the relationship between individual PSI and individual productivity:

Hypothesis 4: Organizational-level PSI moderates the effect of individual-level PSI on individual level productivity. The moderation effect is quadratic such that it is negative when individuals have low PSI and positive when individuals have high PSI

Aside from the direct effects of the organizational-level PSI, there is sometimes disagreement about the nature of an open source community. Peer production efforts are distributed, often with anonymous participants. As a result, they can be chaotic and sometimes anarchistic. There may be clusters of participants who view the community identity one way and clusters who view it a different way. Such divisions can occur in more traditional organizations as well and can be a motivating force for change (Creed, DeJordy, & Lok, 2010). If a group is a vocal minority who has a concentration of power that is due to something other than numbers, they may have an influence on how individuals are treated in the organization regardless of the organization's generally agreed upon identity. In the case where there is high agreement on the organizational-level PSI, the organizational-level PSI will have its full moderating effect on the relationship between individual-level PSI and individual productivity, as per hypothesis 4. However, where there is low agreement, the disagreement in the community can be taken as a measure of chaos that will work opposite to the effects of organizational-level PSI diluting the positive or negative benefit that it has on the effects of

individual PSI on individual productivity. Figure 3 depicts the effect of agreement on the effect of organizational PSI.
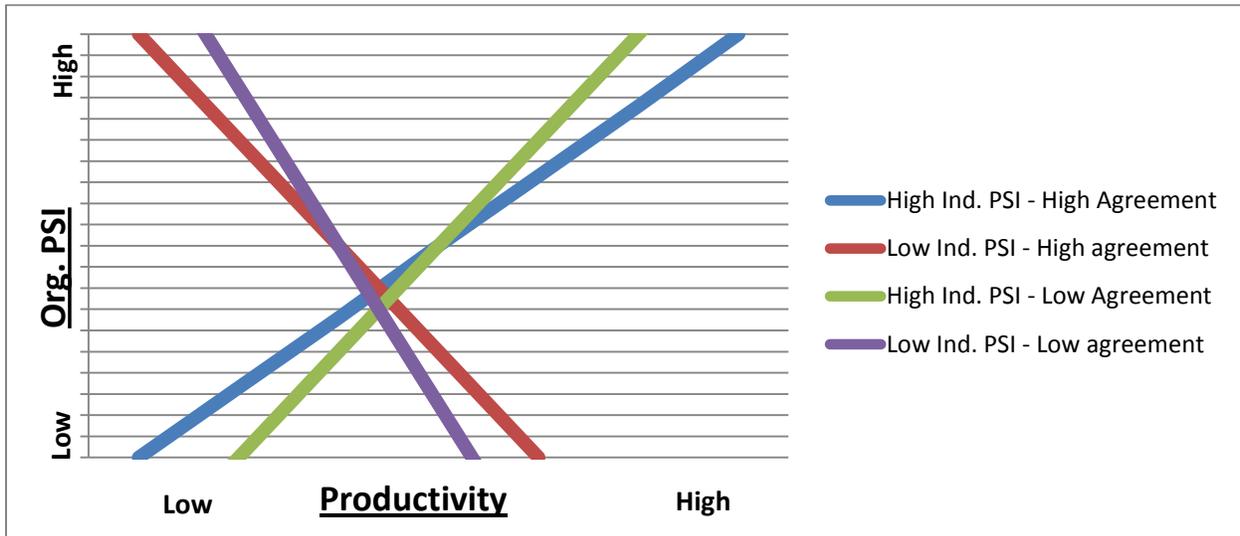

Figure 3: Effect of level of agreement about organizational PSI on the effect of organizational PSI

For example, if an individual with high PSI is embedded in an open source community with high organizational-level PSI, hypothesis 4 suggests that productivity will be at its highest. However, if there is disagreement about the organizational-level PSI, and a small but powerful minority of organization members believe that the organization actually has low PSI, independent of the net effect of the organization's Psi, the disagreement will cause conflict such that the individual will have an identity gap relative to the small but powerful minority who hold the opposing view of the organization's identity. The effect of the level of agreement can best be represented as a positive and linear moderation of the moderation of the organizational-level PSI on the effect of individual-level PSI on individual productivity:

Hypothesis 5: The level of agreement about an organization's PSI positively and linearly moderates the moderating effect of the organization's PSI on the effect of individual-level PSI on the individual's productivity.

Data collection and proposed analysis

The target individual-level population is all participants in open source communities. The target organizational-level population is all open source communities. Given the self-categorization of participation in open source communities, the identification of a sample frame from which to calculate response rates is not possible. In order to get a sample that best represents the target population, random stratified sampling (Singleton & Straits, 2005) is selected as it allows for an oversampling of smaller communities to ensure that the likelihood of higher number of responses from larger communities due to their size does not affect the statistical power of observed effects. In order to conduct appropriate multi-level analysis that balances the power trade-offs of sampling across different levels, a target number of open source communities 30 is selected, with the aim of getting at least 30 individual responses from each individual community (Hofmann, 1997). Open source projects are randomly selected from the list of active projects at Sourceforge.net (Sourceforge.net, 2012), ensuring that at least 10 of the projects are selected from the top 50 projects in terms of activity, 5 of the projects have more than 1000 registered members, 5 of the projects have less than 50 registered members, and that there is a range of project activity and member count for the remaining 10 projects.

A questionnaire is developed as part of a larger research project that includes the measures for this research and is distributed to all the registered members of the community by email via the community's official mailing list with a description of the research outline and an invitation to participate. It will ask individual respondents to provide their responses for

each of the measures and contribution metrics under consideration. In order to capture

members that participate in the community via means other than participation on the active

mailing list or on the Sourceforge.net development platform, questionnaires are also

distributed to attendees at two open source conference in the cities of Toronto and Ottawa,

Canada. The questionnaires distributed at the conferences will also ask respondents to identify

their primary open source community in order to match up their responses to a specific

organization. This method allows for the inclusion of organizations that might not be

represented on Sourceforge.net and allows for snowball referral of informants to strengthen

the number of responses from individual organizations that might not be represented.   The

target number of individual responses is 1000. Collectively, these two methods of data

gathering should provide an adequately diverse sample of both open source communities and

types of individual respondents.

Once the list of open source communities represented in the sample is identified, the

performance data of the community along with the controls, such as size, will be gathered by

visiting the individual open source community official web pages and searching through their

published statistics and documentation. The foundational documents of the each open source

project will also be used to identify the historical roots of the project.

Following the methods described by Bliese (2000), individual responses will first be

assessed for agreement according to the r_{wg} agreement index to determine if they can be

aggregated to represent the individually-reported organizational-level factors (Mathieu &

Taylor, 2007). A direct consensus model about the identity of the organization seems to be the

best way of aggregating the data based on how it is operationalized and collected (Chan, 1998).

The organizational-level factors, both individually-reported and collected will then be assessed

using analysis of variance (ANOVA) and ICC(1) and ICC(2) to determine if there is sufficient

between-group variance for meso-analysis (Mathieu & Taylor, 2007). Within-level and cross-

level theorized relationships will be tested using hierarchical linear models that reflect the

nested nature of the informants within the organizations (Hofmann, 1997). OLS regression is

not appropriate as it cannot account for the correlation between individual responses from

members of the same community, which would confound even within-level effects. The r_{wg}

agreement index will also be used to assess hypothesis 5.

Measures and controls

The individual-level identity measures were adapted from Vesala, Peura, and McElwee's

(2007) self-categorization measures of entrepreneurial identity. As profit seeking identity is

related to but distinct from entrepreneurial identity, the questions in the measure were

adapted to focus on the context of open source community participation and to focus on self-

categorization which is analogous to the identity comparison process and classification process.

The measure uses three questions that ask respondents their level of agreement, on a Likert 1

to 5 scale, with the following statements: "I am someone who participates partly because I

want to make money", "When I first joined the project, I saw myself as able to make money

with it", and "In the future, I would like to be someone who makes money with the project".

The organizational-level identity measures were adapted from a sample of Ohe et al.'s

(1991) corporate difference index, allowing for individual respondent's perceptions of their

open source community's profit seeking identity relative to other open source communities. Once again, the questions were adapted to focus on the profit seeking and participation context, taking into account the possible structural effects that could result from the organizational identity. The measure uses three questions that ask respondents their level of agreement, on a Likert, 1 to 5 scale, with the following statements: "Compared to other communities, this community is more profit-oriented", "In this community, participants who want to make money are welcome", and "The governance structure in this community enables participants to make money".

Individual-level productivity is operationalized as the measures that are tracked for individual contributions to an open source project. The selected measures are the number of bug reports submitted in the past year, the number of years that the individual has been a member of the community, the number of people that the individual has referred to and gotten involved in the community, the number of submissions the individual has made to the project's code base in the past year, and, the number of lines of code the individual has written and contributed to the project's code base in the past year. These measures can be collected by self-report from the individuals and validated against the data gathered by the projects' development platforms.

Organizational-level productivity is operationalized in a way similar to individual-level productivity and is represented by the measures that are tracked for the open source project as a whole as determinants of its health and activity. The selected measures are the number of bug reports submitted to the project in the past year, the number of registered members in the

project, the number of active members relative to total registered members, the rate of increase (decrease) in project membership, the total number of submissions to the project's code base over the past year, and, the total number of lines of code contributed to the project's code base in the past year. Research has suggested that these measures may be good predictors of organizational health and survival (Crowston, Howison, & Annabi, 2006).

Open source communities are diverse so it is necessary to control for heterogeneity between communities that is unrelated to the constructs under investigation in this research. A major confounding factor is project size, represented by the total number of lines of code that make up the project's code base, and, represented by the total number of project members. Larger projects are likely to have significantly different activity effects than smaller projects and the way identity develops and shapes behaviour in small vs. large communities is also likely to be different. Another potential confounding factor that should be controlled is the historical origin of the project and its community. Open source projects that began as the result of a for-profit company releasing a previously proprietary piece of software as open source and encouraging external development have evolved differently and have different identity characteristics that are confounded with profit seeking than open source projects that began as grassroots projects with a few participants and grew into prominence and more wide spread relevance as members formed a community around it. While there are a variety of ways of representing historical origin, a dummy variable that separates corporate historical origin from grassroots historical origin should be sufficient to control for the influence for the purpose of the constructs in question. The foundational documents of the each open source project will be used to identify the historical roots of the project.

Discussion and conclusion

This research makes four contributions. First, it has the potential to affect the productivity levels of open source communities by improving the understanding of the impact of organizational identity on levels of contribution. Rather than focusing exclusively on the individual or the open source project as per previous research, this study considers cross-level effects that may be significant factors for productivity outcomes that could not have been identified with traditional research designs.

Second, the outcomes of this research have the potential of informing the long term debate in open source communities about profit orientation and its impact on the community. Profit seeking in open source communities is still a contentious issue, with many participants arguing that it does more harm than good for the community. By taking an identity perspective, this research allows for a more nuanced understanding of profit seeking at different levels, and grounds the phenomenon in a specific theoretical framing for interpretation and extension to practical considerations.

Third, this research makes a contribution to the identity literature by examining a non-traditional form of invisible identity that I call profit seeking identity. While entrepreneurial identity, professional identity, and social identity in the form of identifiable group membership have all been extensively considered, profit seeking identity is distinct in that it focuses on a motive and participation orientation of an individual or organization that is salient, prototyped, and has outgroup distinctiveness. Its presence or absence as one of the multiple identities that individuals and organizations hold may have an impact on measurable individual and

organizational outcomes that is independent of the more complex and multi-faceted identity measures in the literature. Future research may wish to consider the discriminant validity of profit seeking identity as compared to other identities such as entrepreneurial and professional for predicting individual and organizational productivity measures, particularly to determine if it is a nested or independent identity.

Fourth, this research makes a contribution to the organizational theory literature by examining meso-level identity effects in a novel organizational context. Open source communities are non-traditional organizational forms and the meso-level effects have not previously been considered in the literature. Future research could contrast meso-level identity effects in open source communities to those in traditional organizational forms in the literature. There is also room to extend the research to the institutional level, where effect of the sources of organizational myths, rules, and norms that shape the identities and structure of organizations could be examined for meso-level effects. Of particular interest would be an examination of the separate impact of institutional factors on individual and organizational identity development, along with the individual-organizational identity gap in order to map out how the effects across different levels (institutional-individual and institutional-organizational) might be represented by the measurable individual-organizational identity gap. Such an effect could have profound impact on organizational performance.

References

Albert, S., & Whetten, D.A. 1985. **Organizational identity**. In Cummings, L.L., & Straw, B.M. Eds. Research in organizational behavior. Volume 7: 263-295. Greenwich, CT, USA: JAI Press.

Alvesson, M. 1994. Talking in Organizations: Managing Identity and Impressions in Advertising Agency. *Organization Studies*. 15: 535-563.

Ashforth, B.E., & Mael, F. 1989. Social identity theory and the organization. *Academy of Management Review*. 14(1): 20-39.

Ashforth, B.E., Rogers, K.M., & Corley, K.G. 2011. Identity in organizations: Exploring cross-level dynamics. *Organization Science*. 22(5): 1144-1156.

Bliese, P.D. 2000. **Within-group agreement, non-independence, and reliability – Implications for data aggregation and analysis**. In Klein, K.J., & Kozlowski, S.W.J. Eds. Multilevel theory, research, and methods in organizations. San Francisco, CA: Jossey-Bass.

Bonaccorsi, A., & Rossi, C. 2006. Comparing motivations of individual programmers and firms to take part in the open source movement: from community to business. *Knowledge, Technology and Policy*. 18(4): 40-64.

Brown, A.D., & Starkey, K. 2000. Organizational identity and learning: a psychodynamic perspective. *Academy of Management Review*. 25(1): 102-120.

Chan, D. 1998. Functional relations among constructs in the same content domain at different levels of analysis: A typology of composition models. *Journal of Applied Psychology*. 83(2): 234-246.

Chreim, S., Williams, B.E., & Hinings, C.R. 2007. Interlevel influences on the reconstruction of professional role identity. *Academy of Management Journal*. 50(6): 1515-1539.

Clair, J.A., Beatty, J.E., & Maclean, T.L. 2005. Out of sight but not out of mind: Managing invisible social identities in the workplace. *Academy of Management Review*. 30(1): 78-95.

Corley, K.G., & Gioia, D.A. 2004. Identity ambiguity and change in the wake of a corporate spin-off. *Administrative Science Quarterly*. 49(2): 173-208.

Creed, W.E.D., DeJordy, R., & Lok, J. 2010. Being the change: Resolving institutional contradiction through identity work. *Academy of Management Journal*. 53(6): 1336-1364.

Crowston, K., Howison, J, & Annabi, H. 2006. Information systems success in free and open source software development: theory and measures. **Software Process Improvement and Practice**. 11(2): 123-148.

Demil, B., & Lecocq, X. 2006. Neither market nor hierarchy nor network: the emergence of bazaar governance. **Organization Studies**. 27(10): 1447-1466.

Eclipse Foundation. 2012. **About the Eclipse Foundation**. Available at: http://www.eclipse.org/org/

Ely, R.J. 1994. The effects of organizational demographics and social identity on relationships among professional women. **Administrative Science Quarterly**. 39(2): 203-238.

Fiol, C.M. 2001. Revisiting an identity-based view of sustainable competitive advantage. **Journal of Management**. 27: 691-699.

Foreman, P., & Whetten, D.A. 2002. Members' identification with multiple-identity organizations. **Organization Science**. 13(6): 618-635.

Gioia, D.A., Schultz, M., & Corley, K.G. 2000. Organizational identity, image, and adaptive instability. **Academy of Management Review**. 25(1): 63-81.

Golden, K.A. 1992. The individual and organizational culture: Strategies for action in highly-ordered contexts. **Journal of Management Studies**. 29: 1-21.

Golden-Biddle, K., & Rao, H. 1997. Breaches in the boardroom: organizational identity and conflicts of commitment in a nonprofit organization. **Organization Science**. 8(6): 593-611.

Gutierrez, B., Howard-Grenville, J., & Scully, M.A. 2010. The faithful rise up: Split identification and an unlikely change effort. **Academy of Management Journal**. 53(4): 673-699.

Hofmann, D.A. 1997. An overview of the logic and rationale of hierarchical linear models. **Journal of Management**. 23(6): 723-744.

Hogg, M.A., & Terry, D.J. 2000. Social identity and self-categorization processes in organizational contexts. **Academy of Management Review**. 25(1): 121-140.

Humphreys, M., & Brown, A.D. 2002. Narratives of organizational identity and identification: a case study of hegemony and resistance. **Organization Studies**. 23(3): 421-447.

Ibarra, H. 1995. Race, opportunity, and diversity of social circles in managerial networks. **Academy of Management Journal**. 38(3): 673-703.

Lerner, J., & Tirole, J. 2001. The open source movement: key research questions. **European Economic Review**. 45: 819-826.

Lerner, J., & Tirole, J. 2002. Some Simple Economics of Open Source. *The Journal of Industrial Economics*. 50 (2): 197-234.

Mathieu, J.E., & Taylor, S.R. 2007. A framework for testing meso-mediational relationships in organizational behavior. *Journal of Organizational Behavior*. 28: 141-172.

MERIT. 2006. **Study on the economic impact of open source software on innovation and the competitiveness of the information and communication technologies sector in the EU.** Commissioned Report for the European Commission. Available at: http://ec.europa.eu/enterprise/sectors/ict/files/2006-11-20-flossimpact_en.pdf

Milton, L.P., & Westphal, J.D. 2005. Identity confirmation networks and cooperation in work groups. *Academy of Management Journal*. 48(2): 191-212.

Ohe, T., Honjo, S., Oliva, M., & Macmillan, I.C. 1991. Entrepreneurs in Japan and Silicon Valley: A study of perceived differences. *Journal of Business Venturing*. 6: 135-144.

Peteraf, M., & Shanley, M. 1997. Getting to know you: a theory of strategic group identity. *Strategic Management Journal*. 18: 165-186.

Pratt, M.G., Rockmann, K.W., & Kaufmann, J.B. 2006. Constructing professional identity: The role of work and identity learning cycles in the customization of identity among medical residents. *Academy of Management Journal*. 49(2): 235-262.

Reichers, A.E. 1987. An interactionist perspective on newcomer socialization rates. *Academy of Management Review*. 12(2): 278-287.

Scott, S.G., & Lane, V.R. 2000. A stakeholder approach to organizational identity. *Academy of Management Review*. 25(1): 43-62.

Shepherd, D., & Haynie, J.M. 2009. Birds of a feather don't always flock together: Identity management in entrepreneurship. *Journal of Business Venturing*. 24(4): 316-337.

Singleton, R.A., & Straits, B.C. 2005. **Approaches to social research**. 4th Edition. New York, NY, USA: Oxford University Press.

Sourceforge.net. 2012. **About Sourceforge.net**. Available at: http://sourceforge.net/about

Stallman, R. 1999. The GNU operating system and the free software movement. In DiBona, C., Ockman, S., & Stone, M. Eds. **Open sources: voices from the open source revolution**. Sebastopol, CA: O'Reilly Media Inc.

Vesala, K.M., Peura, J., & McElwee, G. 2007. The split entrepreneurial identity of the farmer. *Journal of Small Business and Enterprise Development*. 14(1): 48-63.

von Hippel, E., & von Krogh, G. 2003. Open Source Software and the "Private-Collective" Innovation Model: Issues for Organization Science. *Organization Science*. 14 (2): 209-223.

von Krogh, G., Spaeth, S., & Lakhani, K.R. 2003. Community, joining, and specialization in open source software innovation: a case study. *Research Policy*. 32: 1217-1241.

Wry, T., Lounsbury, M., & Glynn, M.A. 2011. Legitimating nascent collective identities: coordinating cultural entrepreneurship. *Organization Science*. 22(2): 449-463.