Journal of the Association for Information Systems

Volume 23 | Issue 6 Article 8

2022

When to Signal? Contingencies for Career-Motivated Contributions in Online Collaboration Communities

Jeongsik "Jay" Lee
Drexel University, jaylee@drexel.edu

Hyunwoo Park
Seoul National University, hyunwoopark@snu.ac.kr

Michael Zaggl
Aarhus University, zaggl@mgmt.au.dk

Recommended Citation
Lee, Jeongsik "Jay"; Park, Hyunwoo; and Zaggl, Michael (2022) "When to Signal? Contingencies for Career-Motivated Contributions in Online Collaboration Communities," *Journal of the Association for Information Systems*, 23(6), 1386-1419.
DOI: 10.17705/1jais.00765
Available at: https://aisel.aisnet.org/jais/vol23/iss6/8

Journal of the Association for Information Systems (2022) **23**(6), 1386-1419
doi: 10.17705/1jais.00765

RESEARCH ARTICLE

ISSN 1536-9323

When to Signal? Contingencies for Career-Motivated Contributions in Online Collaboration Communities

Jeongsik "Jay" Lee,[1] Hyunwoo Park,[2] Michael A. Zaggl[3]
[1]LeBow College of Business, Drexel University, U.S.A., jaylee@drexel.edu
[2]Graduate School of Data Science, Seoul National University, South Korea, hyunwoopark@snu.ac.kr
[3]Dept. of Management, School of Business & Social Sciences, Aarhus University, Denmark, zaggl@mgmt.au.dk

Abstract

Online collaboration communities are increasingly taking on new roles beyond knowledge creation and exchange, especially the role of a skill-signaling channel for career-motivated community members. This paper examines the contingency effects of job-market conditions for career-motivated knowledge contributions in online collaboration communities. From the data of individual-level activities in a computer programming-related online Q&A community (Stack Overflow), merged with job-market data for software developers, we find robust evidence of a positive association between community members' career motivations and their knowledge contributions. More importantly, we find that this positive relationship is strengthened by job-market conditions: the number of vacancies in the job market, the expected salaries from these jobs, and the transparency in the flow of career-related information between the community and external recruiters. We contribute to the motivation literature in online collaboration communities by identifying and substantiating the role of contextual factors in mobilizing members' career motivation. Our study thus offers novel insight into how career motivation can be effectively utilized to motivate contributors in these communities. Our findings also point to a possible paradigm change by characterizing online collaboration communities as emerging institutions for career motivation and skill signaling.

Keywords: Community-Market Transparency, Crowdsourcing, Job-Market Signaling, Job Vacancies, Motivation, Online Collaboration, Online Collaboration Communities, Salary, User Contribution

Giri Kumar Tayi was the accepting senior editor. This research article was submitted on September 16, 2020 and underwent two revisions. Michael A. Zaggl is the corresponding author.

Look at the first page or two of Stack Overflow users. Pick anyone at random. Look at three or four of the highly voted answers they wrote. If you've ever hired a programmer in your life, it's obvious those people are all some of the best programmers you could ever hire. Then keep going deeper and deeper. Scroll to page 5. Edit the URL and go right to page 100 where they have reputations in the 3000 range. Look at everyone. With the very rare exception of someone who got a lot of points for a silly answer, these are all obvious superstar programmers... the kind that most teams would kill for.
– Joel Spolsky, co-founder of Stack Exchange[1]

1 Introduction

Knowledge production and sharing in large crowdsourced online collaboration communities (OCCs)—such as Q&A websites (Majchrzak et al., 2021; Zhao et al., 2016), communities of practice (Hara & Hew, 2007; Kudaravalli & Faraj, 2008; Wasko & Faraj, 2005), collaborative knowledge production (Han et al., 2020; Nov, 2007), and open source software development (Ke & Zhang, 2010; von Krogh et al., 2012)—is an increasingly significant

[1] http://programmers.stackexchange.com/questions/20407

phenomenon (Bateman et al., 2011; Bock et al., 2015; Butler et al., 2014; Faraj et al., 2011; Zhao et al., 2016). Finding out what motivates contributions to such crowd-based knowledge production and dissemination is key to enhancing innovation and knowledge flows and thus improving social welfare. One of the most relevant drivers of motivation is the starting or advancing of a professional career related to the community's knowledge or skill domains, which we refer to as *career motivation*. For example, hobbyist product designers can kick off a professional career by demonstrating their talent in design communities (Füller et al., 2007).

Traditionally, people's career motivations and job seeking have been directly embedded in institutions of higher education such as universities. By certifying skills in the form of degrees and diplomas, these institutions allow graduates to signal their job-related skills to potential employers (Holmström, 1999; Spence, 1973). Such signaling is especially required when it is difficult to directly assess the skills of potential employees, as is the case with high-skilled work. We argue that OCCs can serve a role similar to that of traditional institutions of higher education by producing virtual reputations as a certification of skills.

Career motivation has received significant attention in the context of open source software (OSS) development as a driver of community-member contributions (Hann et al., 2013; Ke & Zhang, 2010; Lakhani & Wolf, 2005; Lerner & Tirole, 2002; Roberts, Hann, & Slaughter, 2006; von Krogh et al., 2012). However, the existing literature has implicitly characterized career motivation as an unconditional trait, largely ignoring its contextual influences (Lakhani & Wolf, 2005; Roberts et al., 2006; von Krogh et al., 2012). Contributors to OSS projects are perceived as motivated by career prospects and other forms of motivation, such as ideology (Stewart & Gosain, 2006), reciprocity, or fun (Shah, 2006). In this context, it is suggested that such motivation remains independent of likely varying contextual conditions. We challenge this characterization of career motivation and put forward the following research question:

RQ: How do contextual conditions facilitate or inhibit career motivation in OCCs?

We theorize that the *number of job vacancies*, *expected salary*, and *community-market transparency* are major drivers facilitating career motivation. We turn to the signaling and labor market literature (Flyer, 1997; Freeman, 1975; Holmström, 1999; Siow, 1984; Spence, 1973) to operationalize these facilitators. Although the signaling and labor market literature suggests that these factors are themselves positively related to career motivation (Connelly et al., 2011; Spence, 1973), we know little about their interactions with career motivation in shaping contribution behavior in OCCs. Moreover, while the "crowding out" of intrinsic motivations by extrinsic motivations (Frey & Jegen, 2001; Frey & Oberholzer-Gee, 1997) is relatively well-documented in OCCs (Ke & Zhang, 2010; Zhao et al., 2016), the net effect of career motivation as a form of extrinsic motivation (von Krogh et al., 2012) remains unclear. Because multiple motivations may be simultaneously at play in OCC contexts, career motivation may or may not displace other, perhaps stronger, sources of intrinsic motivation, rendering the net effect on member contributions ambiguous.

Our empirical strategy is to utilize individual-level contribution data from Stack Overflow, an OCC founded in September 2008. In Stack Overflow, community members post questions and answers to computer programming-related issues for a large variety of programming languages. In particular, we exploit Stack Overflow Careers,[2] which is an intracommunity career service. Stack Overflow Careers allows members to post a curriculum vitae (CV) and external recruiters to browse CVs and identify promising job candidates. Because members' achievements in the community—in the form of reputation scores—appear in their CVs, members can signal their programming skills directly to recruiters. Based on how much effort the members exerted to construct their CVs, we quantify their interest in programming-related jobs (i.e., member career motivation). Moreover, the introduction of Stack Overflow Careers was an exogenous event that drastically improved career-related information transparency in the community. To capture the external conditions of the job market for programming, we obtained data from IT Jobs Watch,[3] a company that tracks IT-related job advertisements in the United Kingdom. The data contain the quarterly number of job vacancies and the offered salaries in over 110 programming languages, which allowed us to precisely match the programming-language-specific job conditions to each member's primary programming language and their contributions to the corresponding language domain.

Our study contributes to the growing literature on crowdsourced OCCs. First, we characterize OCCs as emerging, alternative signaling institutions. Our theory and empirical evidence suggest that OCCs could pose a challenge to the current paradigm on institutions of

[2] http://business.stackoverflow.com/careers/

[3] http://www.itjobswatch.co.uk

career signaling since knowledge workers can now signal their quality and job skills without necessarily relying on traditional signaling institutions. We submit that this new paradigm of signaling provides considerable advantages in terms of cost and accuracy. It is also quite likely that OCCs will play a more fundamental role as signaling institutions in the near future. Second, we develop a comprehensive theory on career motivation in OCCs. Our theory contextualizes career motivation by highlighting the enablers of career motivation that are rooted in job-market conditions, thereby emphasizing the strength of extrinsic motivations at play in OCCs. Finally, we extend career motivation beyond the context of OSS development (Ke & Zhang, 2010; Lakhani & Wolf, 2005; Lerner & Tirole, 2002; Roberts et al., 2006) to OCCs, which is arguably a more general form of private-collective mode of knowledge production (von Hippel & von Krogh, 2003).

2 Conceptual and Theoretical Background

2.1 Online Collaboration Communities

New and diverse forms of crowd-based knowledge production and exchange have recently emerged. Examples include online communities for new product development (Füller et al., 2014; Füller et al., 2007; Nambisan, 2002), Q&A communities (Majchrzak et al., 2021; Zhao et al., 2016), communities of practice (Hara & Hew, 2007; Wasko & Faraj, 2005), collaborative knowledge production communities (Nov, 2007), user support forums (Jabr et al., 2014), and open source software communities (Ke & Zhang, 2010; von Krogh et al., 2012). We refer to these forms collectively as online collaboration communities. They are collections of individuals who voluntarily suggest ideas, provide feedback, address questions, and solve others' problems.

All these types of OCCs have mainly been conceptualized and examined around their utility for innovation, problem-solving, and mutual learning (Nambisan, 2002). In this paper, we deviate from this perspective and look at these communities as institutions for career motivation and skill signaling. By emphasizing career motivation, we highlight that OCCs contain diverse benefits for their members and that these benefits can complement one another. While one member might be interested in solving a problem, another one may want to promote her career. We argue that both goals can be reached—and more precisely, one contains the solution to the other—if the members are matched with the right institutional environment.

2.2 Career Motivation and Signaling Theory

From a theoretical standpoint, the argument of career motivation can be built on signaling theory (Connelly et al., 2011; Holmström, 1999; Spence, 1973). Signaling is a mechanism that allows job seekers to overcome the information asymmetry of their individual skills and qualities between them and potential employers. Job seekers reveal their unobservable skills and qualities by making costly investments such as earning a degree. Less-skilled job seekers must exert greater efforts than more-skilled job seekers—compensating for their lack of skills—to achieve the equivalent. Hence, only the more-skilled job seekers choose to signal. As a result, potential employers, as receivers of the signal, are able to distinguish more-skilled job seekers from less-skilled peers. In the context of OCCs, signaling theory represents a mechanism for linking career motivation with contribution behavior. Members contribute to show their otherwise-hidden skills. For example, software developers demonstrate their programming expertise by participating in open source software development projects (Lerner & Tirole, 2002).

However, there is inconsistent empirical evidence of the effectiveness of career motivation for contribution behavior. For instance, some studies document a strong effect of career motivation (Hars & Ou, 2002) while others show a relatively modest effect (Lakhani & Wolf, 2005) or no effect (Ke & Zhang, 2010). One explanation for the inconsistent findings is the "crowding out" between different types of motivation (Frey & Jegen, 2001; Frey & Oberholzer-Gee, 1997). Career motivation, as a strong extrinsic form of motivation, can overshadow intrinsic motivation, thus leading to an overall reduction in contributions. Crowding out has been observed, for example, in the opinions of Swiss citizens; while the majority (50.8%) were willing to support the construction of a nuclear waste disposal site, when financial compensation was offered, support plummeted to less than 25% (Frey & Oberholzer-Gee, 1997, p. 749). Crowding out is particularly relevant in OCC contexts because multiple motivations may collide and their compound effect may simply not add up (Alexy & Leitner, 2011; Roberts et al., 2006). For example, Zhao et al. (2016) found that extrinsic rewards for knowledge sharing undermine members' attitudes toward knowledge sharing.

This points to the need for a contextualization (Johns, 2006; Leidner, 2020) of career motivation. Specifying the factors that moderate the effect of career motivation on the outcome of interest will help reconcile the mixed findings in prior studies. To identify contextual conditions, we build on the human capital theory and labor market theory.

2.3 Contextual Conditions for Career Motivation

Research on occupational decision-making explains how individuals behave to signal effectively. Specifically, two research areas deal with that question: human capital theory and labor market theory. Human capital theory focuses on individuals and their education, which is considered an investment with the possibility of future returns (Becker, 1962, 1964; Schultz, 1962). An individual's career decisions can be seen as a cost-benefit analysis, balancing investments in education and future returns in the form of career benefits (e.g., salaries, social status, and vocational fulfillment) (Boskin, 1974). Human capital theory provides an insightful ground for our characterization of OCCs as signaling institutions. The signaling activities of individuals can be seen as investments in the same way as educational investments. Career-related activities such as signaling incur costs, which, for signaling to be worthwhile, must be more than offset by the expected payoff.

The literature on human capital theory has identified two job-market factors on how individuals balance costs and benefits to inform their occupational decision-making based on the number of job vacancies and expected salary. The *number of job vacancies* reflects the opportunities for an individual to obtain employment on a certain career path. The more vacancies that are open, the higher the likelihood of finding a job given the investment in education (or signaling) (Zarkin, 1985). Thus, the number of job vacancies is an indicator of the probability of successfully achieving career goals. Indeed, empirical research on occupational choice in traditional contexts suggests that market demand have a strong influence on career decision-making (Flyer, 1997; Freeman, 1975; Siow, 1984).

Human capital theory also emphasizes *expected salary* as an influencing factor for occupational choice (Flyer, 1997; Siow, 1984). Salary represents not only a benefit on its own but is also associated with an increase in social status and prestige (Magee & Galinsky, 2008), which are important considerations in career decisions. Thus, the higher the salary offered for the job, the greater the expected returns from finding a job, given the investment in skill signaling. Therefore, the number of job vacancies (i.e., the probability of getting a job) and the expected salary (i.e., the payoff from the job, holding constant the probability of getting a job) together determine the expected payoff from engaging in career-related activities such as signaling via OCC contributions.

Labor market theory, in contrast, focuses on market conditions instead of differences among individual job seekers (Hicks, 1963). That is, job seekers orient their career decisions based on the demand for labor, which is reflected by the number of available positions and the wages offered. Therefore, we find that despite the difference in focus, the labor market literature echoes precisely the same two factors identified from the human capital perspective. In addition, the labor market literature also emphasizes information asymmetries and information flows between job seekers and recruiters (Autor, 2001; Stigler, 1962). In particular, the seamless flow of information—which, in the context of OCCs, we call *community-market transparency*—has been characterized as a critical and desirable criterion for labor markets to be vital (Isgin & Sopher, 2015; Wadensjö, 2015). We elaborate on community-market transparency as a third contingent factor for career-motivated members' contributions. Before proceeding, we note that applying the notion of information transparency from the labor market literature to OCCs is a novel contribution because transparency here connects community members and recruiters *outside* the communities. Beyond the literature on transparency in general labor markets (Autor, 2001; Isgin & Sopher, 2015; Stigler, 1962), there is some research on transparency *inside* communities (Dabbish et al., 2012). However, transparency *between* a community and its external stakeholders has not been studied.

Community-market transparency conceptualizes how information flows between OCC members and external recruiters might influence the contribution behavior of career-motivated members. Analogous to the definition of transparency in labor markets (Autor, 2001; Wadensjö, 2015), we define community-market transparency as *the ability of job seekers and recruiters to observe information about each other*. Hence, greater transparency means a more efficient exchange of information between job seekers and recruiters. In a more transparent job-market environment, job seekers obtain more information about job opportunities and conditions for the same search effort (or the same information for a lower effort), and recruiters learn more about community members, particularly their job skills, at a given cost. Both of these increase the probability of matching between external employers and job seekers (Autor, 2001; Kroft & Pope, 2014). Note that such information about available jobs is highly relevant and useful for job seekers because that information is specifically targeted and broadcast within the community, rather than relying on general information that applies to all job seekers in the broader labor market. Better prospects of job matching might encourage career-motivated members to be more active in signaling through contributions to their community. Further, prompted by greater information availability, other members might realize their latent career motivation and start taking job-market conditions into account in deciding on their contribution level. Greater community-market transparency also decreases the costs of skill signaling.

In contrast, under conditions of low transparency, much of the signaling effort gets lost in the process of information transmission. At least part of this loss must be borne by the community members, reducing their incentives to signal through contributions. Thus, community-market transparency is likely to be highly relevant to the link between career motivation and contribution activity.

3 Hypothesis Development

To elaborate on our theory, we develop four testable hypotheses (see Figure 1). We first develop a baseline hypothesis on the relationship between members' career motivation and their contributions. We then build an argumentation for the three conditions as moderators of this baseline relationship.

3.1 Career Motivation and Knowledge Contributions

Most fundamentally, we expect that a significant share of OCC members are interested in the prospects of entering or advancing a career related to their community activities. This form of motivation has been extensively researched in the context of OSS development (Hann et al., 2013; Ke & Zhang, 2010; Lakhani & Wolf, 2005; Lerner & Tirole, 2002; Roberts et al., 2006) and, albeit to a much lesser extent, in OCCs, specifically in profession-specific communities of practice (Wasko & Faraj, 2005), innovation communities (Seidel & Langner, 2015), and Q&A communities (Xu et al., 2019). Members will, of course, vary in the degree of their career motivation. Some members may be uninterested in jobs but still contribute for other reasons (Roberts et al., 2006), while others might have strong job-seeking inclinations. The latter, we refer to as career-motivated members or job seekers.

The theoretical backbone for linking career motivation to contribution behavior is signaling theory (Connelly et al., 2011; Holmström, 1999; Spence, 1973), which suggests that career-motivated members will try to demonstrate their otherwise unobserved skills by contributing to their community. Here, their ability to provide costly contributions is considered evidence of their skills. Therefore, for a baseline, we expect a positive association between a member's career motivation and their contribution level.

H1a: A community member's career motivation is positively related to the member's contribution level.

Alternative to this expectation of a positive effect, there could be a net reduction of knowledge contributions due to the effect of career motivation. With the strengthening of career motivation as an extrinsic form of motivation, intrinsic motivation might be reduced, causing crowding out to occur (Zhao et al., 2016) and resulting in a net loss of contributions. This crowding out can happen both at the member level and the community level. With career opportunities through community activity becoming more salient, members who have been contributing primarily based on intrinsic motivations such as fun or a sense of community, may now find themselves much less motivated to contribute. At the community level, where some members are primarily driven by career motivation and others are motivated by other factors, the boost in contributions by the former type of member may fail to sufficiently compensate for the reduction in contributions by the latter type. In either case, a net reduction in contributions may result. Such crowding out of intrinsic motivation could also lead to a constant level of contributions if the boost from extrinsic motivation is precisely counterbalanced by the reduction from intrinsic motivation, leading to no change in the contribution level. To account for this possibility, we pose the following competing hypothesis:

H1b: A community member's career motivation is negatively related or unrelated to the member's contribution level.

3.2 The Number of Job Vacancies as a Facilitator

Job attainability, represented by the number of job vacancies, influences the relationship between career motivation and contribution behavior. Career-motivated community members increase their signaling effort in response to greater job availability in their skill domains. In line with signaling theory, they apply a cost-benefit-based rational choice perspective. The more jobs that are available for a specific skill, the more attractive it will be for individual members to make investments in those activities because they may help them obtain a job. Empirical investigations in traditional labor markets support a positive effect of job availability by showing that the market demand (i.e., the number of available jobs) has a positive effect on career decision-making such as job-market entry and educational degree attainment (Flyer, 1997; Freeman, 1975; Zarkin, 1985). This suggests that job vacancies are likely to positively affect the signaling activities of career-motivated members.

This argumentation is also consistent with the expectancy theory of motivation (Vroom, 1964). The theory, rooted in the question of career motivation and job-related motivation (Lynd-Stevenson, 1999; Vroom, 1964), postulates the relationship between an individual's effort and the expected probability of goal achievement. The higher the expected probability of achieving the goal, the more effort individuals will exert (Behling & Starke, 1973).

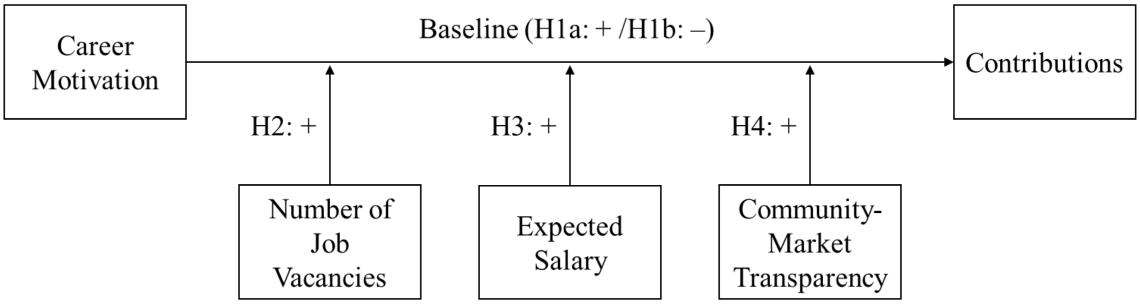

Figure 1. Research Model

Thus, a greater number of job vacancies—representing a higher probability of achieving career goals—increases the signaling effort of career-motivated members. Therefore, we expect the association between a member's career motivation and their contributions to strengthen as the number of job vacancies in their skill domain increases.[4]

H2: The number of job vacancies in a community member's skill domain positively moderates the relationship between the member's career motivation and the member's contribution level.

3.3 Expected Salary as a Facilitator

The expected gains from signaling—i.e., the typical salary of a job—is also a key facilitator of the relationship between career motivation and member contributions. In general, there is a strong positive link between monetary incentives and human behavior (Lazear, 2000). We know from the research on traditional labor markets that income has a strong influence on career selection decisions (Flyer, 1997; Siow, 1984). All else being equal, one would prefer a higher salary over a lower one. As in the number of job vacancies, job seekers can easily observe the salary level expected from the jobs in their skill domain and react to it. Monetary compensation from a given job is a relatively straightforward criterion for any career-motivated member to adopt in making career-oriented decisions such as contribution investments. Specifically, career-motivated members will invest more in contribution-based signaling if the expected salaries from the jobs that can be obtained are higher. Hence, we predict a positive moderation of the expected salary on the baseline relationship between career motivation and member contributions.

Broadly speaking, the rational choice argument has received only partial and indirect empirical support in the literature. Hann et al. (2013) found significant wage differentials between software developers depending on their status in the community, thus supporting the notion of signaling and linking it to salaries. In contrast, Bitzer et al. (2017) failed to find such a relationship. However, our focus here is on the moderating effect of expected salary, rather than its direct impact on contribution-based signaling. Insofar as the expected salary from available jobs functions as a form of monetary incentive and thereby strengthens the signaling motivation of career-motivated members, we can expect career-motivated signaling to increase with the expected salary level. Thus, we hypothesize:

H3: The expected salary from available jobs in a community member's skill domain positively moderates the relationship between the member's career motivation and the member's contribution level.

3.4 Community-Market Transparency as a Facilitator

The third and final factor in our framework is the ease of information flow between the community and the job market (specifically, external recruiters), which we refer to as community-market transparency. In general, transparency reduces the costs of transactions (for empirical evidence in financial markets, see Bessembinder & Maxwell, 2008). Thus, we expect that in the context of OCCs, greater transparency in the job-matching process should reduce frictions and search costs by improving information flows between the supply side (job seekers in the community) and the demand side (external recruiters). Community-market transparency benefits job-seeking members pursuing their career goals in several ways. Under conditions of greater transparency, such members can observe more and better information about opportunities in the job market. Community-market transparency also lowers

[4] Note that motivational crowding out only applies to the baseline relationship (H1a and H1b). Any improvement in the conditions for career motivation, such as the ones we examine in this study, only adds to extrinsically motivated contributions *without* triggering an additional reduction in intrinsically motivated contributions. Such crowding out, if any, is limited to the baseline relationship, and the proposed moderation only concerns the *marginal* effect of this additional facilitation in career motivation on member contributions.

the effective cost of sending signals because it allows members to target specific employers instead of spreading their signaling effort across numerous targets. More-focused signaling efforts on a select set of employers at a given cost allow career-motivated members to produce more and/or higher-quality signals.

On the demand side, community-market transparency has several merits for external recruiters, allowing them to more easily target talent. Greater transparency means improved quality of the signals, which helps recruiters screen job candidates more effectively, thereby facilitating their hiring decisions. These benefits for recruiters, in turn, create positive feedback for the supply side of the market, encouraging more community members to seek career opportunities through the community. Gains from greater market transparency will then draw more recruiters to the market. "Thick market externalities" (Diamond, 1982; Gan & Li, 2016) may thus result—the job market becomes populated with more players on both sides, leading to a higher likelihood of matching between job seekers and recruiters and thereby improving conditions for both sides.

The benefits associated with community-market transparency suggest that career-motivated members will likely adjust their contribution level in response to the degree of transparency. Under conditions of greater transparency, any changes in the job market are transmitted to career-motivated members with less information loss. Likewise, given greater transparency, career-motivated members can expect their signals to reach recruiters more effectively. Hence, career-motivated members will find signaling via contributions to be more attractive for pursuing their career goals and thus increase contributions when community-market transparency is high.

Further support comes from a behavioral perspective as well as the expectancy theory of motivation (Vroom, 1964), as discussed in H2. Greater community-market transparency increases the prospects of fulfilling career goals specifically because career-motivated members will be more convinced that the signals they create in the online community can reach the intended receivers (i.e., external recruiters). Without a clear path to these recruiters, members will be less motivated to send a signal via contributions because there is little guarantee that the signal will reach the intended receivers. Thus, under conditions of low transparency, members' career motivations may remain largely latent and members may be less likely to consider contributions as signaling opportunities. This weakens the link between career motivation and contribution level. However, with a clearer and more effective information channel in place, the community and the activities inside it become more visible to external recruiters. Realizing this will, in turn, boost the motivation of career-motivated members to send signals through increased contributions. As a result, the link between career motivation and contribution level is strengthened. That is, higher levels of career motivation justify greater levels of contribution. Together, these reasons suggest that under conditions of greater community-market transparency, the association between career motivation and contribution level becomes stronger.

H4: Information transparency between the supply and demand sides of the job market (i.e., community-market transparency) positively moderates the relationship between a community member's career motivation and the member's contribution level.

4 Empirical Design

4.1 Research Setting and Data Construction

Our research context is Stack Overflow, a free and public online Q&A platform for computer programming-related issues (see Appendix A for details). Career motivation is highly relevant to the members of this OCC. According to a survey of over 56,000 members,[5] more than three quarters of the respondents expressed interest in a programming-related career by indicating that they were either actively looking for a job (15.5%) or open to new job opportunities (62.7%).

Stack Overflow has a career service, "Stack Overflow Careers."[6] This service allows members to create a CV certifying their reputation scores. Paying recruiters can custom-search the CVs and list job advertisements. Thus, members can leverage their activities and reputation in the community as a signal and thus present themselves to potential employers. The community introduced this service in two stages. The initial version (launched on November 3, 2009)[7] charged fees to both community members (a one-time fee of $25) and recruiters (amount undisclosed). In the subsequent version "Stack Overflow Careers 2.0" (launched on February 23, 2011),[8] Stack Overflow removed the fees for community members.[9] This

[5] http://stackoverflow.com/insights/survey/2016

[6] This service was later renamed "Jobs" (http://stackoverflow.com/jobs?med=site-ui&ref=jobs-tab).

[7] http://blog.codinghorror.com/stack-overflow-careers-amplifying-your-awesome/

[8] https://blog.stackoverflow.com/2011/02/careers-2-0-launches/

[9] Instead of fees, an invitation from a peer member was required to sign up to Stack Overflow Careers (we thank an anonymous reviewer for pointing us to this). Evidence

change has considerably relaxed the members' constraints, making it much easier for them to use Stack Overflow Careers.

Critically, we believe that the introduction of Stack Overflow Careers has decidedly increased community-market transparency. Stack Overflow's original intention was to help members find better jobs[10] rather than to directly promote contributions. With the career service platform in place, job-seeking members can much more efficiently signal their quality to recruiters[11]. In particular, they can directly target the employers and jobs that match their career goals. Likewise, recruiters can identify and access qualified candidates more easily and at considerably lower costs. Hence, we assume that the introduction of Stack Overflow Careers significantly enhanced the transparency of the job market relevant to the members in the OCC.

We acquired the entire member activity data in Stack Overflow over five years, from the outset of the community to June 2013. The raw dataset contains 14,630,209 Q&As written by 2,055,496 unique members. We also obtained job-market demand data from IT Jobs Watch, a company that maintains historical records of the IT job market in the United Kingdom. The dataset contains quarterly data on the number of job vacancies and average salary offers for 113 programming languages from 2006 to 2015. Thus, each of these programming languages represents a segment of the IT job market. We matched the two data sources following the process elaborated in Appendix B and constructed a panel dataset.[12] For a meaningful construction of the various measures employed in our analysis, we limited the sample to members who contributed a total of at least 10 answers. Our results are robust to variations in this minimum threshold. The panel has 684,000 member-quarter observations of 72,444 members. Figure B1 in the Appendix describes the sample construction.

4.2 Variables

4.2.1 Dependent Variables

Our dependent variable should capture the member's contribution level. Members can contribute to the community by either posting frequently, making well-received contributions, or both. However, posts can be well received by others in the community only when a member makes a post. Hence, we primarily measure member contributions by the number of posts. To construct the *quantity of contributions*, we counted the number of total posts that a member made in each quarter.[13]

We also considered *vote-weighted contributions* to measure how others in the community perceive contributions. Other community members can evaluate contributions by voting up (+1) or down (-1). We constructed *vote-weighted contributions* using the sum of net votes for the contributions in each member-quarter.

For robustness checks and post hoc analyses (see Appendix C), we constructed two auxiliary dependent variables that were broken down by contribution type. The first is the number of questions and the number of answers counted separately for each member-quarter (i.e., the total number of posts are broken down into questions and answers). The second one breaks down the counts by contribution domain: primary programming-language domain, non-primary language domain, and non-language domain. For each domain, we counted the number of posts for each member-quarter.

4.2.2 Independent Variables

For the baseline relationship, we needed a measure that captures the degree to which a member is interested in seeking a career. For this, we exploited the CVs posted on Stack Overflow Careers, which reveal two important pieces of information about members' career

suggests that invitations did not specifically target career-motivated members, had no specific qualifications, were virtually costless to make, and did not constitute any endorsement to the invited. Hence, we have no reason to suspect that the invitation-based system has introduced a systematic bias by, for instance, overrepresenting career-motivated members.

[10] "Many of you reading this have great jobs that you love, but many more do not, and don't realize that Stack Overflow has a product that could help get them a much better job" (http://meta.stackoverflow.com/questions/310066/stack-overflow-serving-programmers-even-better).

[11] Even without this channel, the members could still document their community activities in CVs and use them to signal their quality to target employers. This is theoretically possible and may have been happening in practice even before the launch of the service. However, this method of signaling is inefficient for members because there is no standard format for collating relevant activities and achievements. It is also ineffective because there is no guarantee that the target employer would pay attention to this particular piece of information. From an employer's standpoint, this information is difficult to verify because it lacks any endorsement or certification by the community. Absent the career service, it is costly for the recruiters to screen community members and identify those who possess the skills in demand.

[12] Note that we address potential issues arising from matching global-level data (Stack Overflow) with UK-level data (IT Jobs Watch) in our robustness checks later.

[13] The contribution data are more fine-grained, but our time unit of sample is a quarter in order to match the frequency of job-market data.

motivation: whether a member has posted a CV to the site and what content the CV contains. We used the estimated level of effort to create CV as a proxy for the member's career motivation. All else being equal, a longer CV takes more time and effort to compose and thus represents a greater level of career motivation than a shorter CV. This increased effort indicates that members posting CVs are likely to be more inclined to take a professional stance on their OCC work. They may thus be more engaged with career opportunities. To construct the CV-based measure of career motivation, we first searched the CVs of all members in our sample. Of these, 34,007 members (46%) posted their CVs. We then constructed *career motivation* by counting the number of words contained in each CV. We assigned zero for the members with no CV posted. For post hoc analyses, we also constructed category indicators of career motivation by classifying the members into three groups: *no CV*, *short CV*, and *long CV*. We used these category variables to represent different degrees of career motivation in a post hoc analysis (see Appendix A). As detailed below, we added controls capturing other member attributes such as experience and writing style to ensure that CV length indeed measures the career motivation of the community members rather than their differences in other dimensions.

To test H2, we captured the *number of job vacancies* as the quarterly count of job advertisements for each programming language, compiled by the company IT Jobs Watch. For H3, we computed *expected salary* as the quarterly average of salaries offered for permanent jobs in a given programming language, also collected by IT Jobs Watch. Because a member may contribute to multiple language domains but our job-market data are specific to a programming language, for each member-quarter, we assigned the primary language based on the member's activity in the preceding four quarters.[14]

For the construction of *community-market transparency* (H4), we exploited the introduction of Stack Overflow Careers. As discussed earlier, this career service promotes the information flow in both directions between community members and external recruiters. Recruiters can advertise job vacancies in a standardized format, which makes members aware of the demand and makes finding relevant job postings easier. The service also improves the quality of members' signaling of their skills. Reputation scores earned through knowledge contributions are available to recruiters, who can then compare across multiple candidates on a common standard. Stack Overflow Careers initially charged fees for both the members and the recruiters but later made it free for members after some time. We chose to distinguish between before and after the elimination of member fees ("Stack Overflow Careers 2.0") and constructed a binary variable called *community-market transparency* (or simply *transparency*) to indicate low (0) and high (1) transparency between the community and external job markets.

It is important to note that the three variables representing the contingent factors in our framework are exogenous to the community members. The number of vacancies and expected salaries are beyond the control of individual members and difficult for them to predict in a precise manner. The elimination of fees in Stack Overflow Careers was also exogenous.

4.2.3 Control Variables

Most of our controls consist of a series of fixed effects to account for possible confounding influences from other covariates of contribution activity. To account for time-invariant member-level unobserved heterogeneity, we included dummies for individual members. Also, because other job-market conditions and contribution behavior may differ across programming languages, we included dummies for programming languages. Quarter dummies were included to account for possible temporal variations. Programming language-fixed effects and quarter-fixed effects were included in all estimation models. However, testing the baseline (H1a and H1b) prevents us from including member-fixed effects because our measure of career motivation is time-invariant within the member. Hence, when testing the baseline, we used a random-effects model.

Members may vary in experience, which may confound the effect of career motivation. Thus, we controlled for member experience by including community tenure, measured by the number of days since the member posted the first time in the community.[15] To account for the possible heterogeneity in writing propensity across members, we included the average length of posts (time-invariant) for each member in random-effects models. This variable may also capture some of the heterogeneity across members in writing style (some members tend to write longer than others do) and hence partially correct for baseline intermember differences in CV length, our measure of career motivation.

[14] Though allowed to change from quarter to quarter, the primary language of a member was quite stable across time. In the quarter-by-quarter transition matrix of primary language, the occurrence-weighted average of the diagonal elements was 87.9% (i.e., 87.9% of the time, the primary language stayed the same as that of the previous quarter). On average, members exhibited 1.62 primary languages over the span of their activity.

[15] Ideally, we would have liked to use the number of years (or months) of programming experience as identified from the CV, but that would result in missing observations for all members who did not post their CVs.

4.2.4 Descriptive Statistics of Variables

Table 1 shows the summary statistics of the variables used in the analysis.[16] The member-quarter panel in our sample is unbalanced and the variables were summarized as raw values before log transformation was applied. During our sample period, members contributed an average of 10.8 posts per quarter and received about 2.4 votes per post. The vast majority of these posts are answers (9.0; 83.2%) to other members' questions, while questions represent a much smaller portion (16.8%). About 64% of the posts (6.9) are related to at least one of the programming languages. The average career motivation (i.e., the count of words in posted CVs) is 183. On average, 11,107 permanent jobs were available each quarter and the average salary for these jobs was £44,943. About 72% of the member-quarter observations occurred under conditions of high community-market transparency (i.e., after the introduction of Stack Overflow Careers 2.0).

Looking at the pairwise correlations between variables (Table 2), no pair exhibited a correlation high enough to cause a multicollinearity concern, except for the correlation between transparency and quarter as a continuous variable (r = 0.83). This is natural, given the construction, but it is not an issue for identification because time effects were absorbed through quarter dummies. Many of the variables were highly skewed. Hence, we log-transformed all continuous variables.

4.3 Estimation

We estimated the following log-linear model, with all the continuous variables on both sides of the equation taking the natural logs of the corresponding value:

$$q_{i,t} = \alpha + \beta C_i + \sum_{z=1}^{3} y_z M_{z,t} + \sum_{z=1}^{3} \delta_z (C_i \times M_{z,t}) + \eta_{i,} + \lambda_j + \xi_t + \varepsilon_{i,t} \quad (1)$$

where $q_{i,t}$ denotes either the number of posts (quantity) or the vote-weighted posts, C_i indicates career motivation, $M_{z,t}$ represents one of the moderators (*number of job vacancies*, *expected salary*, and *community-market transparency*), η_i are member-fixed effects (omitted in random-effects models), λ_j are programming-language-fixed effects, ξ_t quarter-fixed effects, and $\varepsilon_{i,t}$ an idiosyncratic error term. Here, β is the coefficient of interest in the baseline (H1a and H1b), and δ_1 through δ_3 are the coefficients of interest in H2-H4.

We estimated the equation using both random-effects and fixed-effects panel OLS regression models. The random-effects model is primarily used to estimate β, which tests the baseline. Note that β cannot be estimated in the fixed-effects model because C_i is time-invariant and thus absorbed by member-fixed effects, η_i. Note also that *community-market transparency* is a dummy variable based on calendar time and thus its coefficient is not separately identified (absorbed by one of the quarter-fixed effects, ξ_t).

5 Results

5.1 Hypothesis Tests

Model 1 in Table 3 shows the relationship of career motivation with the quantity of contributions using the (log) number of posts as the dependent variable. The coefficient on *career motivation* is strongly positive (Model 1), supporting H1a and rejecting the notion of crowding out (H1b). The estimated elasticity (0.032) implies that doubling the length of a posted member CV is associated with a 3.2% increase in the member's volume of contributions to the community.[17]

Models 2-6 include member-fixed effects, where the direct effect of *career motivation* is not identified due to collinearity. In Model 3, we find a significant and positive interaction between *career motivation* and *number of job vacancies* showing that the career motivation-contribution relationship is positively moderated by job availability. This suggests that career-driven members contribute more frequently when there are more job vacancies in their programming language available in the market, which supports H2.

Model 4 tests the interaction between *career motivation* and *expected salary*. This interaction is significant and positively related to contribution quantity, indicating that career-driven members respond to a greater degree when the available jobs offer higher salaries. The first-order effect of *expected salary* is statistically insignificant, as in the other models. This suggests that the effect of expected salary on contributions occurs mainly in conjunction with the member's career motivation. Hence, we found support for H3.

[16] Due to missing values in some variables, the final sample is slightly reduced to 682,710 member-quarter observations of 72,427 members.

[17] This may not seem large in terms of magnitude, but note that this is an average across all community members in the sample and, as shown later, masks considerable heterogeneity across members.

Table 1. Summary Statistics

Variable	N	Mean	Std. Dev.	Min	Max
Quantity of contributions (Q + A)	682,710	10.82	34.99	0	2004
Vote-weighted contributions	682,710	23.98	125.67	-52	21406
# Questions	682,710	1.78	5.28	0	333
# Answers	682,710	9.04	34.01	0	2004
# Posts on programming languages	682,710	6.90	25.31	0	1765
Career motivation	279,227	183.48	431.18	4	5491
Number of job vacancies	682,710	11107.14	7765.16	1	28877
Expected salary	682,710	44943.03	5429.44	19000	100000
Community-market transparency	682,710	0.72	0.45	0	1
# Days since joining	682,710	674.09	407.89	45	1749
Average length of posts	682,710	737.80	318.80	108.07	5204.58
Quarter	682,710	48.08	4.53	36	54
Note: Quarter variable is serialized in reference to Q1 2000 as Quarter 1.					

Table 2. Pairwise Correlations

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
(1) Quantity of contributions											
(2) Vote-weighted contributions	0.79*										
(3) # Questions	0.26*	0.16*									
(4) # Answers	0.99*	0.79*	0.11*								
(5) # Posts on programming lang.	0.95*	0.77*	0.22*	0.94*							
(6) Career motivation	0.13*	0.12*	0.03*	0.13*	0.11*						
(7) Number of job vacancies	0.01*	-0.02*	0.01*	0.01*	0.00*	0.03*					
(8) Expected salary	-0.00*	0.01*	-0.03*	0.00	0.01*	-0.02*	0.05*				
(9) Community-market transparency	-0.05*	-0.10*	-0.05*	-0.04*	-0.04*	-0.04*	0.11*	0.00			
(10) # Days since joining	-0.10*	-0.08*	-0.12*	-0.08*	-0.09*	0.09*	0.08*	0.06*	0.43*		
(11) Average length of posts	0.01*	0.02*	0.12*	-0.00*	0.02*	0.06*	0.02*	0.05*	0.04*	-0.02*	
(12) Quarter	-0.06*	-0.13*	-0.07*	-0.05*	-0.05*	-0.05*	0.09*	0.02*	0.83*	0.52*	0.04*

Table 3. Quantity of Contributions

	(1)	(2)	(3)	(4)	(5)	(6)
(Log) # of days since joining	-0.356** (0.004)	-0.284** (0.007)	-0.284** (0.007)	-0.284** (0.007)	-0.284** (0.007)	-0.284** (0.007)
(Log) average length of posts	0.352** (0.008)					
(Log) career motivation	0.032** (0.002)					
(Log) # job vacancies	0.027* (0.011)	0.027* (0.011)	0.021+ (0.011)	0.027* (0.011)	0.027* (0.011)	0.023* (0.011)
(Log) expected salary	0.060 (0.044)	0.025 (0.045)	0.025 (0.045)	-0.036 (0.048)	0.024 (0.045)	-0.028 (0.048)
(Log) career motivation × (Log) # job vacancies			0.004** (0.001)			0.003** (0.001)
(Log) career motivation × (Log) expected salary				0.048** (0.014)		0.041** (0.014)
(Log) career motivation × Transparency					0.013** (0.003)	0.012** (0.003)
Constant	0.610 (0.508)	3.046** (0.523)	3.053** (0.524)	3.053** (0.522)	3.059** (0.523)	3.069** (0.524)
Member FE	No	Yes	Yes	Yes	Yes	Yes
Language FE	Yes	Yes	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
N	682710	682710	682710	682710	682710	682710
Within R^2	0.102	0.103	0.103	0.103	0.103	0.103
Note: The dependent variable in all models is the (log) number of posts made by the member in a given quarter. Robust standard errors, clustered by members, are in parentheses. +, *, ** denotes statistical significance at 10%, 5%, and 1%, respectively.						

In Model 5, we obtain a significantly positive coefficient of the interaction term between *career motivation* and *community-market transparency*, consistent with H4. Career-motivated members contribute more frequently when information flows with the job market are under conditions of greater transparency. Considering all explanatory variables at once did not change the results (Model 6).

So far, we have focused on the number of posts (*quantity of contribution*) as a measure of contributions. The other measure would be the *vote-weighted quantity of contributions*, which represents the usefulness of the contributions indicated by votes from other members. From the perspective of an OCC as a knowledge repository, promoting content highly useful to community members is perhaps equally, if not more, important than simple quantity. Posting high-quality content is thus an effective way of signaling skills to potential recruiters. Hence, we looked at vote-weighted contributions by using the (log) number of votes received as the dependent variable (Table 4).

Models 1 through 6 in Table 4 replicate the corresponding models in Table 3. The results are highly similar to that in the analysis of contribution volume: the number of votes received is strongly and positively associated with *career motivation* (Model 1); and the interactions of *career motivation* with the *number of job vacancies* (Model 3), *expected salary* (Model 4), and *community-market transparency* (Model 5) are all significantly positive. Hence, the results suggest that career-motivated members respond to market conditions by adjusting not only the frequency of their contributions but also the perceived quality of those posts. This is reasonable because while the volume of contributions influences the visibility of a member in the community, it is the quality of contributions as perceived by other members—hence reflected in the "reputation" score—that determines the member's status. Hence, members with more career motivation have a clear incentive to boost both the quantity and perceived quality of contributions. Overall, our analysis provides solid support for all hypotheses.

While the results of the vote-weighted contribution analysis are mostly consistent with those of the quantity analysis, two differences are noteworthy. First, across all models in Table 4, the first-order effect of the *number of job vacancies* on vote-weighted contributions was consistently negative, which stands in contrast to the positive effect on contribution quantity (Table 3). This hints at a baseline trade-off between contribution quantity and vote-weighted contributions with increases in job vacancies, such that more frequent posting comes at the cost of the perceived quality of the posts. However, this trade-off seems largely limited to the members who have not posted their CV (hence have zero value on the *career motivation* variable), because the baseline effect estimate on *career motivation* (Model 1) suggests that career-motivated members (those who have posted their CV and hence have positive values for the variable) react positively to both the quantity and perceived quality of contributions. Second, *expected salary*, which does not indicate any direct effect on contribution quantity, generally shows a positive relationship, albeit spotty, with vote-weighted contributions. This suggests that better income prospects in programming jobs generally improve the perceived quality of posts even for the no-CV members, but not the volume of contributions. The boost for career-motivated members is thus above and beyond this baseline effect of expected salary in the market. These differences in the first-order effects of job vacancy and expected salary appear well in line with our theoretical assumptions about how such job-market conditions may influence member behavior.

Having established support for our main hypotheses, in the following we report a series of additional tests to ensure the robustness of our results. Because our empirical strategy builds on behavioral data, we seek to address ensuing issues such as endogeneity concerns and the validity of variables as proxies for theoretical constructs.

5.2 Endogeneity Tests

In the previous analyses, we included member-fixed effects as well as member experience and the average length of posts to control for possible confounds of the effect of career motivation proxied by CV length. Despite these controls, there are reasons to suspect that our results may be subject to endogeneity. For instance, some unobservable quality of members may drive both CV length and contribution level, potentially inflating the estimated effect of career motivation. Although this is not an issue for the tests of H2-H4, where we only exploit the within-member variation using fixed-effects models, it could be an issue for the baseline relationship. To address this concern and to improve causal inference, we employed two approaches that are widely used for handling this type of situation: the coarsened-exact matching (CEM) and instrumental variable (IV) methods.

For the matching analysis, we created a CEM sample between members with a CV and members without a CV based on three member-level variables: primary programming language, (log) number of total posts, and (log) number of total votes received. Because CVs are time-invariant in our data, we used their values aggregated over the sample period for the matching. We enforced exact matching for the primary programming language while coarsening for numbers of posts and received votes using Scott's rule (Scott, 1979).

Table 4. Vote-Weighted Contribution

	(1)	(2)	(3)	(4)	(5)	(6)
(Log) # of days since joining	-0.312** (0.005)	-0.271** (0.009)	-0.271** (0.009)	-0.271** (0.009)	-0.271** (0.009)	-0.271** (0.009)
(Log) average length of posts	0.439** (0.010)					
(Log) career motivation	0.034** (0.002)					
(Log) # job vacancies	-0.052** (0.014)	-0.053** (0.015)	-0.060** (0.015)	-0.053** (0.015)	-0.053** (0.015)	-0.058** (0.015)
(Log) expected salary	0.144*	0.154**	0.154**	0.088	0.152**	0.097
	(0.056)	(0.057)	(0.057)	(0.061)	(0.057)	(0.061)
(Log) career motivation × (Log) # job vacancies			0.005** (0.002)			0.004* (0.002)
(Log) career motivation × (Log) expected salary				0.052** (0.019)		0.044* (0.019)
(Log) career motivation × Transparency					0.014** (0.004)	0.012** (0.004)
Constant	0.586	3.234**	3.241**	3.242**	3.248**	3.259**
	(0.635)	(0.656)	(0.658)	(0.655)	(0.656)	(0.657)
Member FE	No	Yes	Yes	Yes	Yes	Yes
Language FE	Yes	Yes	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
N	678139	678139	678139	678139	678139	678139
Within R^2	0.157	0.157	0.157	0.157	0.157	0.157

Note: The dependent variable in all models is the (log) number of votes received by the member in a given quarter. Robust standard errors, clustered by members, are in parentheses. +, *, ** denotes statistical significance at 10%, 5%, and 1%, respectively.

For the implementation, we used the CEM algorithm in Stata 15 (Blackwell et al., 2009). Appendix D provides details of the matching process. On this matched sample, we reestimated all models with the weights computed by the CEM algorithm. For H1a and H1b, however, we ran a pooled OLS with CEM weights because random-effects panel models do not allow for weights. The results, reported in Models 1 through 6 in Table 5, are largely consistent with those from our main analysis. In fact, the positive effect of *career motivation* became much stronger, with the magnitudes of effect roughly doubling from those of Tables 3 and 4. Given the carefully curated sampling design, this further supports the validity of CV length as a proxy for career motivation. The moderating effects of *job vacancies* and *expected salary* also became stronger, though the magnitude of boosts was smaller. The CEM results do not provide support for H4.

For the IV analysis, we constructed a dummy variable indicating whether a CV contained a link to a member profile at GitHub (a popular public repository of open source codes among others) and used it as an instrument for *career motivation* (i.e., CV length). Our logic here is that the members who listed the GitHub link are likely to be more interested in promoting their programming skills. By design, including a GitHub link will also increase the length of that CV. More importantly, insofar as CV length represents the level of a member's career interest, as we assume, GitHub links will positively correlate with CV length. However, listing a GitHub link itself is unlikely to increase the contribution level, except through its influence on CV length. Thus, this variable is likely to satisfy the exclusion restriction condition for an instrument. 18.8% of the CV members had a GitHub link. Using this instrumental variable, we conducted a two-stage least square estimation of the models. Note, however, that we could only estimate the random-effects model to test H1 because CV length, the variable under scrutiny for being endogenous, is constant within the member and thus the first-stage model cannot be identified with member-fixed effects. Models 7 and 8 in Table 5 show the second-stage results of this IV analysis. For both *quantity* and *vote-weighted contributions*, *career motivation* (instrumented) was a strong positive predictor, with effect magnitudes of more than twice the main results (in Tables 3 and 4). The diagnostics testing underidentification and weak identification both support the validity of the instrument. Taken together, the results from the CEM estimation and the IV estimation suggest that our main results are robust to potential endogeneity in the hypothesized relationships between variables.

5.3 Robustness Tests and Extensions

We performed several additional tests by varying some of the particulars of our empirical design. All the results are detailed in Appendix C and ensure the robustness of our main results while providing more evidence of the validity of our empirical strategy.

Table 5. CEM and IV Regression Results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	CEM Quantity of contributions			CEM Vote-weighted contributions			IV Quantity	IV Vote-weighted
(Log) # of days since joining	-0.347**	-0.317**	-0.317**	-0.294**	-0.305**	-0.305**	-0.364**	-0.299**
	(0.005)	(0.008)	(0.008)	(0.006)	(0.009)	(0.009)	(0.005)	(0.005)
(Log) average length of posts	0.307**			0.388**			0.288**	0.390**
	(0.010)			(0.011)			(0.009)	-0.011
(Log) career motivation	0.058**			0.071**			0.064**	0.081**
	(0.002)			(0.003)			(0.003)	-0.004
(Log) # job vacancies	0.068**	0.035*	0.028+	-0.015	-0.037+	-0.045*	0.035**	-0.052**
	(0.018)	(0.016)	(0.016)	(0.023)	(0.020)	(0.021)	(0.012)	(0.017)
(Log) expected salary	0.036	-0.007	-0.076	0.051	0.110	0.046	0.097+	0.134*
	(0.076)	(0.065)	(0.070)	(0.095)	(0.081)	(0.087)	(0.052)	(0.067)
(Log) career motivation			0.005**			0.007**		
× (Log) # job vacancies			(0.002)			(0.002)		
(Log) career motivation			0.050**			0.046*		
× (Log) expected salary			(0.017)			(0.021)		
(Log) career motivation			-0.001			-0.008*		
× Transparency			(0.003)			(0.004)		
Constant	0.719	3.195**	3.224**	0.992	3.201**	3.231**	---	---
	(0.815)	(0.726)	(0.726)	(1.006)	(0.927)	(0.926)	---	---
Member FE	No	Yes	Yes	No	Yes	Yes	No	No
Language FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	574314	574314	574314	570435	570435	570435	682707	678136
Within R^2	0.073	0.093	0.093	0.098	0.138	0.138	0.042	0.026
Underidentification test	---	---	---	---	---	---	4478.32	4466.98
Weak identification test	---	---	---	---	---	---	26886.97	26881.81

Note: Models 1-6 are estimated with CEM weights as regression weights. Models 7-8 are estimated with instrumental variable regression with the dummy variable on whether a member has one's GitHub profile included in the CV as the instrument. The dependent variable is the (log) number of posts (Models 1-3, & 7) or the (log) number of votes (Models 4-6, & 8) for the member in a given quarter. The underidentification test statistic reports Kleibergen-Paap *rk* LM statistic, and weak identification test statistic reports Kleibergen-Paap rk Wald *F* statistic. The first-stage *F*-statistics (Models 7 & 8) is 5477.47, which is above the conventional threshold of 10 (Stock & Yogo, 2005). Robust standard errors, clustered by members, are in parentheses. +, *, ** denotes statistical significance at 10%, 5%, and 1%, respectively.

These additional tests also helped to add more nuances to our main findings. Specifically, we conducted four sets of analyses. The first one distinguishes between the types of contributions by separately looking at questions posted by a member and answers the member provided to others' questions. This analysis was intended to examine whether the members who we consider to be career motivated optimize their signaling behavior by focusing on contribution activities that might yield more signaling gains (see Appendix C).[18] Second, we divided member contributions into three categories: posts in any programming language, posts in non-programming-language domains, and posts in the members' primary programming-language domain. This was used to check whether career-motivated members concentrated their efforts on specific knowledge domains (i.e., programming languages) that might be more representative of their skills in order to better signal their skills to external recruiters (see Appendix C). Third, we utilized a few variations in the sample by restricting to a subset of members: members who posted their CV, members who joined before the introduction of the career service to the community, and members residing in the EU region. These subsample analyses were meant to ensure that our results hold for a more refined and stringent sample (see Appendix C). Finally, we divided members into three groups by the level of career motivation (proxied by CV length: no career motivation, low career motivation, and high career motivation) and replicated the analysis using these categorical variables instead of the continuous value. This allowed us to examine possible nonlinearity in the effect of career motivation (see Appendix C). Results from these robustness checks were largely consistent with those from our main analysis and support the validity of our empirical design.

[18] In Stack Overflow, both answers and questions receive the same reputation score when voted up, but answers can earn an additional reputation score if they are marked as "accepted" by the members who posted the questions. Also, from the signaling standpoint, contributing answers would likely seem better to external recruiters than asking questions.

6 Discussion

Our results provide strong evidence for our theory on the contextual conditions of career motivation as significant drivers of contributions in OCCs. Using panel analyses of granular member activity data from Stack Overflow merged with programming-language-level job-market data from IT Jobs Watch, we document systematic relationships between members' career interests, job-market demand indicators (number of job vacancies and expected salary), transparency in information flow, and the magnitude of member contributions to the community. We first confirm a net positive link between a member's career motivation and the member's contributions (Lakhani & Wolf, 2005; Roberts et al., 2006; Xu et al., 2019). Our analyses further show that changes in the number of job vacancies and expected salaries for a member-specific programming-language skill positively moderate this career motivation-contribution relationship. Finally, we find that the increased community-market transparency, owing to Stack Overflow Careers, strengthens the positive relationship between career motivation and contribution level.

6.1 Contribution to Theory

Our study offers two major contributions to our theoretical understanding of OCCs. First, our work highlights how the role of OCCs expands beyond knowledge creation and exchange to institutions of skill signaling. Second, we extend the information systems literature on motivation in OCCs by contextualizing the effect of career motivation. Besides these two major contributions, we generalize career motivation beyond OSS development to the broader context of OCCs.

6.1.1 Online Collaboration Communities as Signaling Institutions

Our first major contribution is expanding the role of OCCs by characterizing them as a novel institution for skill signaling and career-seeking. By building on signaling theory (Arrow, 1973; Holmström, 1999; Spence, 1973) and labor market literature (Flyer, 1997; Freeman, 1975), we draw a direct comparison between OCCs and traditional institutions of higher education such as universities. These institutions have historically fulfilled the function of signaling by certifying their members' (i.e., students') achievements and thereby improving their career prospects (Arrow, 1973; Spence, 1973). However, recent research has found that formal academic credentials play a relatively minor role in labor markets, as the majority of employers place greater emphasis on the "job readiness" of candidates (Brown & Souto-Otero, 2020), and there are major discrepancies between job-market needs and educational content (Börner et al., 2018). This is the critical gap that OCCs as non-traditional institutions can fill. They can provide job seekers with the opportunity to signal their immediately applicable skills, and hence job readiness, to external recruiters.

OCCs even have two systematic advantages over traditional signaling institutions of higher education. First, traditional signaling institutions contribute solely to the creation of the signal and the immediate outcomes of the signal production process (i.e., exams and assignments) are usually wasted: "Students do work hard, because of reputation effects, even though it is entirely wasteful from a social point of view" (Holmström, 1999, p. 177). In stark contrast, in OCCs the creation of the signal must provide an immediate benefit for other members. Each contribution has at least one beneficiary (i.e., the asker of the question) and most often multiple beneficiaries, whose votes indicate the value they gain from the contributed content. In this way, the contribution's actual value is tightly coupled with the signal. Because of this tight coupling, we contend that OCCs as institutions of career signaling can be more productive than traditional institutions of signaling—to use Holmström's (1999) language, OCCs reduce the "waste" relative to traditional signaling institutions. Therefore, it might be that, at least for certain job-related qualities, OCCs can substitute for traditional institutions in their role as signal enablers.

Another advantage over traditional signaling institutions is that OCCs can overcome the so-called *career progression paradox*—the problem that recruiters demand experienced job seekers but in order to acquire experience, the job seekers need a job (O'Mahony & Bechky, 2006). By assuming the role of a signaling institution, OCCs reduce this problem because the communities can certify job seekers' skills in the form of digital reputation even *before* they enter the job market. Job seekers can gradually build up their reputation and this process of reputation development does not require a job.

The characterization of OCCs as institutions for career motivation and signaling casts a new light on the functional role of OCCs. Whereas the information systems and management literature has considered OCCs mainly as places of knowledge creation and exchange (Bock et al., 2015; Lou et al., 2013; Majchrzak et al., 2021), they also can be viewed as job-seeking venues. Moreover, there appears a synergy between these two roles. The example of Stack Overflow shows how knowledge creation and exchange are stimulated by the need for signaling and, in turn, how signaling is enabled through knowledge creation and exchange activities within the community.

6.1.2 Contextualization of Career Motivation

The second major contribution of our study lies in contextualizing the effect of career motivation, thereby extending the current theoretical understanding of motivation theory in the information systems and management literature (Johns, 2006). We theorize and empirically corroborate the role of contextual conditions for career motivation in shaping contribution behavior. We ground these contingent factors in the established literature on labor markets and human capital (Flyer, 1997; Freeman, 1975). Our study is one of the first to consider these factors as important conditions that influence the extent to which career motivation translates into tangible contributions to OCCs.

Our contingency view on career motivation in OCCs is thus distinct from the extant literature that has treated career motivation as largely given or as isolated from job-market conditions (Ke & Zhang, 2010; Lakhani & Wolf, 2005; Roberts et al., 2006; Xu et al., 2019). Our study argues for and offers evidence that counters such implicit assumptions. Incorporating contextual conditions into the discussion of career motivation allows for a more complete and coherent picture of career motivation as a contribution driver in these communities. Moreover, in our theory and empirical tests, we focused on the changes in job-market conditions rather than their stock. This allowed us to get at the "sensitivity" of career motivation in its relation to contribution behavior, which adds an important nuance to the scholarly dialogue on this issue.

Our theory contrasts with what we would expect from crowding-out theory (Frey & Oberholzer-Gee, 1997; Zhao et al., 2016) and the research on OCCs that has put forward arguments largely building on intrinsic motivations (von Krogh, 2012). From this perspective, introducing any extrinsic form of motivation might replace intrinsic motivation and result in lower levels of contribution (see H1b). However, we find that this is not the case and a positive impact on contribution behavior is entirely possible. Thus, our findings show that the extrinsic form of career motivation plays a positive role and is further reinforced by job-market-based contextual conditions.

In addition to the contingent effect of job-market parameters, we highlight the importance of community-market transparency. Information transparency has been emphasized by the traditional labor market literature (Autor, 2001; Kroft & Pope, 2014). By also demonstrating the relevance of transparency conditions to community-mediated knowledge production and exchange, we introduce a multi-element contingency theory of career motivation in OCCs.

6.1.3 Generalizing Career Motivation beyond the Context of Open Source Software Development

Besides these two main contributions, our study generalizes the role of career motivation for contribution behavior beyond the context of OSS development, from which the original theory and literature of career motivation emerged (Ke & Zhang, 2010; Lerner & Tirole, 2002; Roberts et al., 2006). In contrast to OSS development, which takes place in relatively more integrated organizations and handles more interrelated tasks such as coding, testing, and code integration (Lee & Cole, 2003), our context is based on flat and loose structures (Faraj et al., 2011) with virtually fully decomposable, atomic tasks. Nonetheless, our study confirms the role of career motivation under these seemingly unfavorable conditions for pursuing career opportunities. Because OCCs provide an arguably more conservative setting, we believe that our findings can be extended to the field of OSS development and support the argument for contribution motivation from a career standpoint (Lerner & Tirole, 2002).

6.2 Managerial Implications

Our study offers important implications for practitioners, especially managers and system designers of OCCs. One of the most critical challenges for OCCs is to attract and retain active members (Bateman et al., 2011; Butler, 2001; Faraj et al., 2011; Ren et al., 2012) in order to continue growing and stay relevant in the increasingly crowded online space. Our study suggests that an effective way of achieving this goal is to improve the information transparency between the community and the external audience such as job recruiters. In light of our findings, OCC designers may want to consider implementing direct channels to potential employers such as Stack Overflow Careers while making them more accessible to their community members, which, as we demonstrate, can help boost contributions in their communities. Of course, different community members have different motivations to participate. However, insofar as some members are motivated by their career interests, though potentially to varying degrees, instituting features that closely connect job markets with the community and improve the information flow between the two sides—which we call community-market transparency—appears to be a very useful strategy for keeping those members motivated and attached to the community.

Enhancing community-market transparency is, of course, possible only if there are significant external career opportunities that demand the knowledge and skills relevant to the community. For instance, for general Q&A communities such as Answers.com and

Quora, it seems difficult to define what the relevant job markets are; in fact, external job opportunities specifically targeting these community members may hardly exist. Hence, community-market transparency is less likely to be relevant to these communities. However, there is a broad range of OCCs that can utilize community-market transparency as a design feature that fosters member participation. For example, many companies build on input from user communities, such as firms relying on community-based customer support (Jabr et al., 2014) or those using ideas and knowledge produced in the community (Afuah & Tucci, 2012; Chen et al., 2011). These types of communities may well benefit from career-motivated participation and be able to create the right conditions to induce career-based motivation. In particular, the hosting companies themselves represent the demand side of the job market. It is thus crucial for the hosting company to utilize this hiring strategy to offer well-paying jobs (i.e., number of vacancies and expected salary, in our framework). More importantly, they must ensure that these career opportunities are directly linked to contributions within the community and that this link is made transparent to both community members and external recruiters (i.e., community-market transparency). Community-market transparency—specifically, transparency inside the OCC—may create potential competition among OCC members. The increased ability of community members to observe each other's activities may spur contributions out of a collective spirit, but it may also engender unhealthy competitive rivalry among members. Thus, designers of OCCs need to develop an elaborate understanding of the information flows inside the OCC.

More broadly, our study also aptly informs the managers of online labor markets such as Upwork, which are increasing in popularity. These online labor markets utilize various reputational mechanisms that, by nature, resemble those used in OCCs, in that workers in these markets earn their reputations based on others' assessment of their task performance. This performance-based reputation has potential beyond its function inside an online labor market because it can be used for signaling capabilities for certain jobs outside that market. For instance, managers of online labor markets could collaborate with external companies that look for candidates for permanent positions. In this case, the reputation that a worker accumulated in the online labor market can function as a useful signal to these external companies seeking to fill a position. Such a connection to a permanent job market would, in turn, likely spur the effort of these workers in their online jobs because it would make their reputation more critical by granting it significance both in online and offline labor markets. This strategy could help improve the overall quality of work inside the online labor market and could thereby create a positive feedback loop between online-offline job markets. However, this advice needs to be taken with caution. More transparent and tighter links to offline job markets may promote migration, potentially reducing the online labor market workforce; however, to compensate for this problem, online job-market managers could, for instance, use flexible pricing schemes for companies seeking job candidates.

On a more general level, our study supports the idea of actively exploiting extrinsic motivations to attract and retain OCC contributors. Despite the possibility of crowding out intrinsic motivations, we show that, at least in our setting, career motivation as an extrinsic form of motivation more than offsets such crowding-out effects, thereby increasing overall contributions. OCC managers are thus well advised to consider exploring various ways to more fully utilize extrinsic forms of motivation—career motivation, in particular—for a more robust design and functioning of their communities.

6.3 Limitations and Future Research Directions

Our study has several limitations, which also provide interesting avenues for future research. First, since we did not know the exact timing of CV posting, our measure of career motivation—CV length—is time-invariant and hence conceals potential dynamics in the effect over time. Thus, regardless of when the members posted their CVs to Stack Overflow Careers, we treated them as if they were career-motivated from the time they appeared in the dataset. However, it is entirely likely that members may have posted their CVs only after they became interested in job seeking, which could have been a long while after they first became active in the community. It is also possible that after posting their CVs, some members may have lost interest in job seeking but still left their CVs on the site. In either case, the actual link between CV length and contribution activity for these members is likely weaker than assumed. Therefore, our estimate based on the time-invariant CV measure is likely conservative. To settle this issue, however, investigations with time-variant data are called for, which would require matching contribution data with a member survey.

Second, our empirical strategy is not suitable for capturing the potential interactions of career motivation with other sources of motivation, especially those of an intrinsic nature. We are not alone in facing this limitation—scholars have generally analyzed career motivation in isolation from other forms of motivation (for an exception, see Roberts et al. 2006). Thus, further research incorporating different motivations and their interactions is needed.

Third, our job-market data cover only the United Kingdom. Ideally, we would have liked to have region-specific data on jobs and salaries to match with members in corresponding regions. Unfortunately, such data at the level of precision of our data are not available. Although we found generally robust results on a narrower sample of members from the EU, caution may be necessary when applying the findings to regions with IT labor markets that deviate considerably from that of the United Kingdom.

7 Conclusion

With the growing popularity of online crowdsourcing communities, knowledge production and dissemination are increasingly reliant on contributions from the crowd. As argued and documented in this study, career motivation fuels voluntary knowledge contributions in these communities, thereby enabling the creation of a "public knowledge good." Through these contributions, community members can signal their job-related skills while also meeting others' needs for problem-solving. Extrinsic incentives from job-market conditions—such as open job vacancies and expected salaries—and information transparency in the community-market interface can help unleash this career motivation, thus amplifying the career-motivated knowledge contributions in these communities. By offering appealing avenues for pursuing career motivations, OCCs may be gradually taking over a function that, up to now, has been exclusively filled by institutions of higher education. Our study has only begun to explore this emerging yet potentially paradigm-shifting phenomenon.

Acknowledgments

Authors are in alphabetical order—all authors are first authors and contributed equally. We thank John Grant from IT Jobs Watch for providing us with the job market data. We also thank the anonymous reviewers and the senior editor, Giri Tayi, for their valuable feedback. We are also grateful to the organizers and reviewers of the Academy of Management Annual Meeting 2017 where an earlier version of this paper received the Best Paper Award in the CTO (OCIS) division. The research of Hyunwoo Park was supported by the National Research Foundation of Korea funded by the Korea government (MSIT) (No. NRF-2022R1C1C1011888).

References

Afuah, A., & Tucci, C. L. (2012). Crowdsourcing as a solution to distant search. *Academy of Management Review*, *37*(3), 355-375.

Alexy, O., & Leitner, M. (2011). A fistful of dollars: Are financial rewards a suitable management practice for distributed models of innovation? *European Management Review*, *8*(3), 165-185.

Arrow, K. J. (1973). Higher education as a filter. *Journal of Public Economics*, *2*(3), 193-216.

Autor, D. H. (2001). Wiring the labor market. *Journal of Economic Perspectives*, *15*(1), 25-40.

Bateman, P. J., Gray, P. H., & Butler, B. S. (2011). Research note: The impact of community commitment on participation in online communities. *Information Systems Research*, *22*(4), 841-854.

Becker, G. S. (1962). Investment in human capital: A theoretical analysis. *Journal of Political Economy*, *70*(5), 9-49.

Becker, G. S. (1964). *Human capital: A theoretical and empirical analysis, with special reference to education*. University of Chicago Press.

Behling, O., & Starke, F. A. (1973). The postulates of expectancy theory. *Academy of Management Journal*, *16*(3), 373-388.

Bessembinder, H., & Maxwell, W. (2008). Markets: Transparency and the corporate bond market. *Journal of Economic Perspectives*, *22*(2), 217-234.

Bitzer, J., Geishecker, I., & Schröder, P. J. H. (2017). Is there a wage premium for volunteer OSS engagement? Signalling, learning and noise. *Applied Economics*, *49*(14), 1379-1394.

Blackwell, M., Iacus, S., Iacus, S., & Porro, G. (2009). cem: Coarsened exact matching in Stata. *The Stata Journal*, *9*(4), 524-546.

Bock, G.-W., Ahuja, M. K., Suh, A., & Yap, L. X. (2015). Sustainability of a virtual community: integrating individual and structural dynamics. *Journal of the Association for Information Systems*, *16*(6), 418-447.

Börner, K., Scrivner, O., Gallant, M., Ma, S., Liu, X., Chewning, K., Wu, L., et al. (2018). Skill discrepancies between research, education, and jobs reveal the critical need to supply soft skills for the data economy. *Proceedings of the National Academy of Sciences*, *115*(50), 12630-12637.

Boskin, M. J. (1974). A conditional logit model of occupational choice. *Journal of Political Economy*, *82*(2), 389-398.

Brown, P., & Souto-Otero, M. (2020). The end of the credential society? An analysis of the relationship between education and the labour market using big data. *Journal of Education Policy*, *35*(1), 95-118.

Butler, B. S. (2001). Membership size, communication activity, and sustainability: a resource-based model of online social structures. *Information Systems Research*, *12*(4), 346-362.

Butler, B. S., Bateman, P. J., Gray, P. H., & Diamant, E. I. (2014). An attraction-selection-attrition theory of online community size and resilience. *MIS Quarterly*, *38*(3), 699-728.

Chen, J., Xu, H., & Whinston, A. B. (2011). Moderated online communities and quality of user-generated content. *Journal of Management Information Systems*, *28*(2), 237-268.

Connelly, B. L., Certo, S. T., Ireland, R. D., & Reutzel, C. R. (2011). Signaling theory: A review and assessment. *Journal of Management*, *37*(1), 39-67.

Dabbish, L., Stuart, C., Tsay, J., & Herbsleb, J. (2012). Social coding in GitHub: Transparency and collaboration in an open software repository. *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work.*

Diamond, P. A. (1982). Aggregate demand management in search equilibrium. *Journal of Political Economy*, *90*(5), 881-894.

Faraj, S., Jarvenpaa, S. L., & Majchrzak, A. (2011). Knowledge collaboration in online communities. *Organization Science*, *22*(5), 1224-1239.

Flyer, F. A. (1997). The influence of higher moments of earnings distributions on career decisions. *Journal of Labor Economics*, *15*(4), 689-713.

Freeman, R. B. (1975). Supply and salary adjustments to the changing science manpower market: Physics, 1948-1973. *American Economic Review*, *65*(1), 27-39.

Frey, B. S., & Jegen, R. (2001). Motivation crowding theory. *Journal of Economic Surveys*, *15*(5), 589-611.

Frey, B. S., & Oberholzer-Gee, F. (1997). The Cost of Price incentives: An empirical analysis of motivation crowding-out. *The American Economic Review*, *87*(4), 746-755.

Füller, J., Hutter, K., Hautz, J., & Matzler, K. (2014). User roles and contributions in innovationcontest communities. *Journal of Management Information Systems*, *31*(1), 273-308.

Füller, J., Jawecki, G., & Mühlbacher, H. (2007). Innovation creation by online basketball communities. *Journal of Business Research*, *60*(1), 60-71.

Gan, L., & Li, Q. (2016). Efficiency of thin and thick markets. *Journal of Econometrics*, *192*(1), 40-54.

Han, Y., Ozturk, P., & Nickerson, J. (2020). Leveraging the wisdom of the crowd to address societal challenges: Revisiting the knowledge reuse for innovation process through analytics. *Journal of the Association for Information Systems*, *21*(5), 1128-1152.

Hann, I.-H., Roberts, J. A., & Slaughter, S. A. (2013). All are not equal: An examination of the economic returns to different forms of participation in open source software communities. *Information Systems Research*, *24*(3), 520-538.

Hara, N., & Hew, K. F. (2007). Knowledge-sharing in an online community of health-care professionals. *Information Technology & People*, *20*(3), 235-261.

Hars, A., & Ou, S. (2002). Working for free? Motivations for participating in open-source projects. *International Journal of Electronic Commerce*, *6*(3), 25-39.

Hicks, J. R. (1963). *The theory of wages* (2nd ed.). Macmillan.

Holmström, B. (1999). Managerial incentive problems: A dynamic perspective. *Review of Economic Studies*, *66*(1), 169-182.

Isgin, E., & Sopher, B. (2015). Information transparency, fairness and labor market efficiency. *Journal of Behavioral and Experimental Economics*, *58*, 33-39. .005

Jabr, W., Mookerjee, R., Tan, Y., & Mookerjee, V. S. (2014). Leveraging philanthropic behavior for customer support: The case of user support forums. *MIS Quarterly*, *38*(1), 187-208.

Johns, G. (2006). The essential impact of context on organizational behavior. *Academy of Management Review*, *31*(2), 386-408.

Ke, W., & Zhang, P. (2010). The effects of extrinsic motivations and satisfaction in open source software development. *Journal of the Association for Information Systems*, *11*(12), 784-808.

Kroft, K., & Pope, D. G. (2014). Does online search crowd out traditional search and improve matching efficiency? Evidence from Craigslist. *Journal of Labor Economics*, *32*(2), 259-303.

Kudaravalli, S., & Faraj, S. (2008). The structure of collaboration in electronic networks. *Journal of the Association for Information Systems*, *9*(10), 706-726.

Lakhani, K. R., & Wolf, R. G. (2005). Why hackers do what they do: Understanding motivation and effort in free/open source software projects. In J. Feller, B. Fitzgerald, S. Hissam, & K. R. Lakhani (Eds.), *Perspectives on free and open source software* (pp. 3-21). The MIT Press.

Lazear, E. P. (2000). Performance pay and productivity. *American Economic Review*, *90*(5), 1346-1361.

Lee, G. K., & Cole, R. E. (2003). From a firm-based to a community-based model of knowledge creation: The case of the Linux Kernel development. *Organization Science*, *14*(6), 633-649.

Leidner, D. E. (2020). What's in a contribution? *Journal of the Association for Information Systems*, *21*(1), 238-245.

Lerner, J., & Tirole, J. (2002). Some simple economics of open source. *Journal of Industrial Economics*, *50*(2), 197-234.

Lou, J., Fang, Y., Lim, K. H., & Peng, J. Z. (2013). Contributing high quantity and quality knowledge to online Q&A communities. *Journal of the Association for Information Science and Technology*, *64*(2), 356-371.

Lynd-Stevenson, R. M. (1999). Expectancy-value theory and predicting future employment status in the young unemployed. *Journal of Occupational and Organizational Psychology*, *72*(1), 101-106.

Magee, J. C., & Galinsky, A. D. (2008). Social hierarchy: The self-reinforcing nature of power and status. *Academy of Management Annals*, *2*(1), 351-398.

Majchrzak, A., Malhotra, A., & Zaggl, M. A. (2021). How open crowds self-organize. *Academy of Management Discoveries*, *7*(1), 104-129.

Nambisan, S. (2002). Designing virtual customer environments for new product development: Toward a theory. *Academy of Management Review*, *27*(3), 392-413.

Nov, O. (2007). What motivates Wikipedians? *Communications of the ACM*, *50*(11), 60-64.

O'Mahony, S., & Bechky, B. A. (2006). Stretchwork: Managing the career progression paradox in external labor markets. *Academy of Management Journal*, *49*(5), 918-941.

Ren, Y., Harper, F. M., Drenner, S., Terveen, L., Kiesler, S., Riedl, J., & Kraut, R. E. (2012). Building member attachment in online communities: Applying theories of group identity and interpersonal bonds. *MIS Quarterly*, *36*(3), 841-864.

Roberts, J. A., Hann, I.-H., & Slaughter, S. A. (2006). Understanding the motivations, participation, and performance of open source software developers: A longitudinal study of the apache projects. *Management Science*, *52*(7), 984-999.

Schultz, T. W. (1962). Reflections on investment in man. *Journal of Political Economy*, *70*(5), 1-8.

Scott, D. W. (1979). On optimal and data-based histograms. *Biometrika*, *66*(3), 605-610.

Seidel, V. P., & Langner, B. (2015). Using an online community for vehicle design: Project variety and motivations to participate. *Industrial and Corporate Change*, *24*(3), 635-653.

Shah, S. K. (2006). Motivation, governance, and the viability of hybrid forms in open source software development. *Management Science*, *52*(7), 1000-1014.

Siow, A. (1984). Occupational choice under uncertainty. *Econometrica*, *52*(3), 631-645.

Spence, M. (1973). Job market signaling. *Quarterly Journal of Economics*, *87*(3), 355-374.

Stewart, K. J., & Gosain, S. (2006). The impact of ideology on effectiveness in open source software development teams. *MIS Quarterly*, *30*(2), 291-314.

Stigler, G. J. (1962). Information in the labor market. *Journal of Political Economy*, *70*(5, Part 2), 94-105.

Stock, J. H., & Yogo, M. (2005). Testing for weak instruments in linear IV regression. In D. W. K. Andrews & J. H. Stock (Eds.), *Identification and inference for econometric models* (pp. 80-108). Cambridge University Press.

von Hippel, E. A., & 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. (2012). How does social software change knowledge management? Toward a strategic research agenda. *Journal of Strategic Information Systems*, *21*(2), 154-164.

von Krogh, G., Haefliger, S., Spaeth, S., & Wallin, M. W. (2012). Carrots and rainbows: Motivation and social practice in open source software development. *MIS Quarterly*, *36*(2), 649-676.

Vroom, V. H. (1964). *Work and motivation*. Wiley.

Wadensjö, E. (2015). Labor market transparency. In J. Forssbaeck & L. Oxelheim (Eds.), *The Oxford handbook of economic and institutional transparency* (pp. 241-257). Oxford University Press.

Wasko, M. M., & Faraj, S. (2005). Why should I share? Examining social capital and knowledge contribution in electronic networks of practice. *MIS Quarterly*, *29*(1), 35-57.

Xu, L., Nian, T., & Cabral, L. (2019). What makes geeks tick? A study of Stack Overflow Careers. *Management Science*, *66*(2), 587-604.

Zarkin, G. A. (1985). Occupational choice: An application to the market for public school teachers. *The Quarterly Journal of Economics*, *100*(2), 409.

Zhao, L., Detlor, B., & Connelly, C. E. (2016). Sharing knowledge in social Q&A sites: The unintended consequences of extrinsic motivation. *Journal of Management Information Systems*, *33*(1), 70-100.

Appendix A: Detailed Research Context Description

Our research context is Stack Overflow, a free and public online Q&A platform. Stack Overflow was launched in September 2008 as a small niche website for hosting Q&As primarily related to computer programming. It has since grown into one of the most popular websites for programmers with over 50 million monthly visitors and cumulative funding of $70 million from institutional investors as of December 2018.[19] Indeed, Stack Overflow touts itself as "the largest online community for programmers to learn, share their knowledge, and advance their careers" (community website). Reflecting its success, the community's business model has spawned more than 170 Q&A sites dedicated to niche topics ranging from bicycles to photography to 3D printing, all of which are housed under the Stack Exchange Network.

Stack Overflow offers an excellent setting to examine the questions we pose in this study. Most importantly, there is an explicit job market that demands knowledge and skills directly relevant to the core activities of the community. Moreover, the community has, over time, introduced an interface through which the members can signal their job-related skills to potential employers, and recruiters can use the interface to screen and identify promising job candidates. Finally, the detailed categorization of topic areas by the community managers allows us to precisely match the skill sets of individual members to the demand in the external job market.

Given the community's broad coverage of various computer-related issues, there can be multiple ways of defining a job market relevant to the community. However, because the community has from its inception most extensively dedicated to providing answers to questions about programming and coding, we focus on the market for software developers as the relevant job market. The market encompasses job demands for a variety of programming languages, allowing us to exploit the variation across different language domains.

[19] http://stackoverflow.com/company/about (Last accessed on December 13, 2018).

Appendix B: Data Matching

We acquired from Stack Overflow, under the Creative Commons BY-SA 3.0 License, the entire activity data spanning 5 years from the outset of the community to June 2013. The raw dataset contains 14,630,209 Q&As written by 2,055,496 unique community members. Each question belongs to up to five categories denoted by "tags" that represent the topics, concepts, or languages the question is directly related to. More than 30,000 tags are covered in the dataset. We also obtained job market demand data from IT Jobs Watch, a UK-based company that collects and maintains records of the number of job vacancies (both permanent and by-contract) in programming languages and the average salary offers (or time rates) for these jobs. The dataset contains quarterly data for 113 programming languages from 2006 to 2015. Each of the 113 programming languages represents a segment of the job market for software developers.

A nontrivial challenge in combining both datasets is to match across different taxonomies of programming languages. Stack Overflow uses highly granular tags to denote the relevant knowledge areas, while IT Jobs Watch tracks the job market demand by programming language. Consequently, a particular programming language in IT Jobs Watch data may be associated with multiple tags according to Stack Overflow's taxonomy. To merge the two datasets, we manually created a concordance between Stack Overflow tags and IT Jobs Watch programming languages. We started with creating a text file that listed all of the over 30,000 tags extracted from the Stack Overflow data. We sorted these tags in descending order of the number of belonging posts so that we pay more attention to popular tags and avoid accidentally missing them. Then, for each programming language defined by IT Jobs Watch, we searched through the list to flag the tags that contain keywords directly related to the language. We iterated this process to refine the mapping. Table B1 shows the final mapping table. In total, 96 programming languages were matched with at least one tag. This language-tag concordance served as the basis for linking member contributions and job market demand (i.e., job vacancies and salaries for a particular programming language).

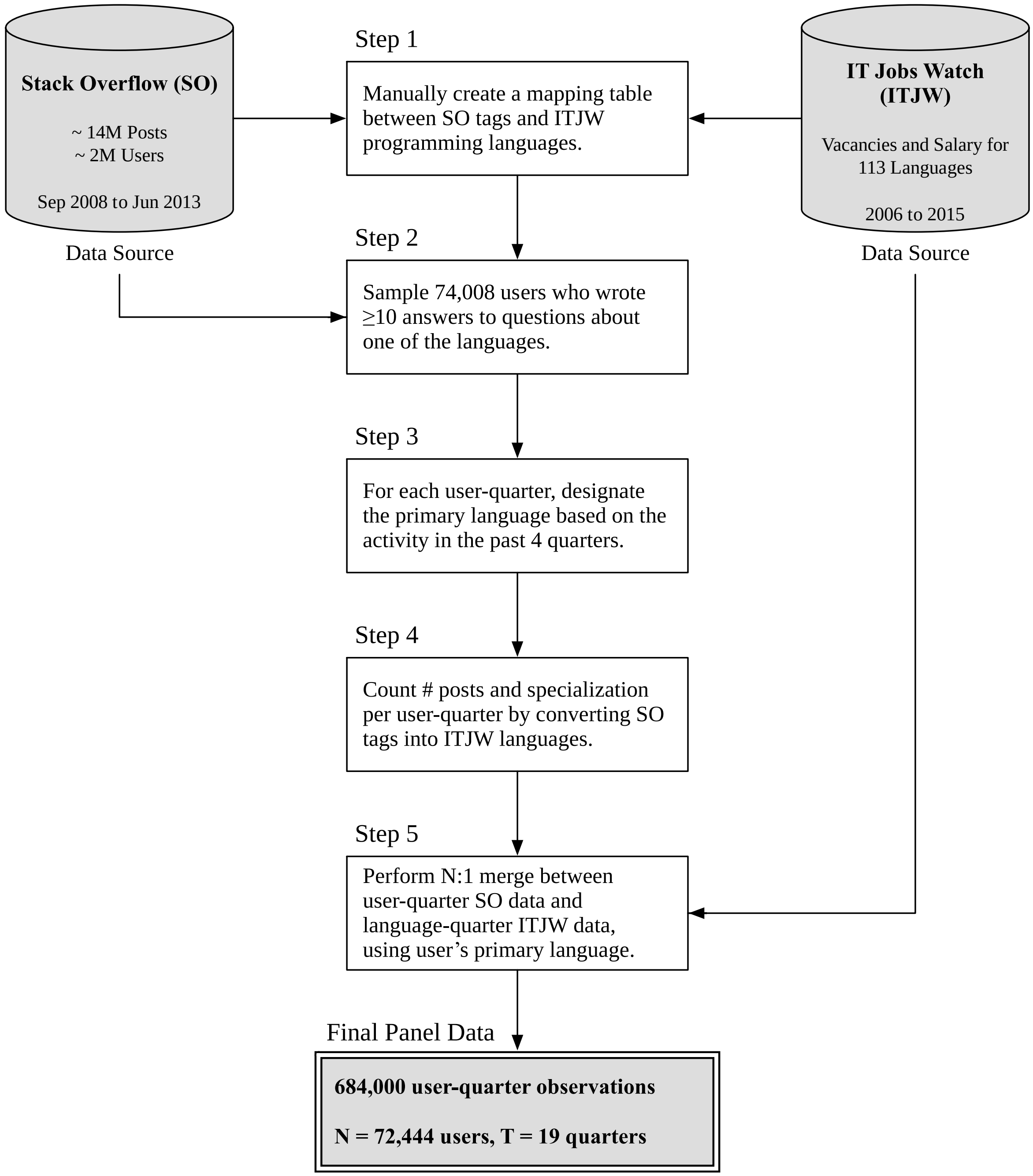

Figure B1. Sampling Process Leading to Final Panel Data

Table B1. Mapping Table between ITJW Languages and Stack Overflow Tags

	ITJW Language	SO Tags			ITJW Language	SO Tags			ITJW Language	SO Tags
1	ABAP	abap (449)		41	GLSL	glsl (1628)		81	RPG	rpg (71)
2	ActionScript	actionscript (6979); actionscript-1 (9)		42	Go	go (2187)		82	RPG III	
3	ActionScript 2.0	actionscript-2 (1539)		43	Groovy	groovy (6191)		83	RPG IV	
4	ActionScript 3.0	actionscript-3 (26912)		44	Haskell	haskell (11384)		84	RPG/400	
5	Ada	ada (507)		45	HiveQL	hiveql (92)		85	Ruby	ruby (74088)

6	ANSI C			46	HLSL	hlsl (463)		86	RubyMotion	rubymotion (243)
7	ANSI SQL	ansi-sql (46)		47	HQL	hql (1788)		87	SAPscript	sap (954); sapscript (2)
8	Apex Code	apex-code (850)		48	IronPython	ironpython (1141); ironpython-studio (21)		88	SAS Macro	sas (1196)
9	Apple Swift	swift (58)		49	J#	j# (54)		89	Scala	scala (16077)
10	AppleScript	applescript (2409)		50	Java	java (426765)		90	sed	sed (4740)
11	AspectJ	aspectj (952)		51	JavaScript	javascript (385054)		91	Shell Script	
12	Assembly Language	assembly (8957)		52	JCL	jcl (95)		92	Smalltalk	smalltalk (639)
13	AWK	awk (4437)		53	Jython	jython (1058); jython-2.5 (44)		93	SOQL	soql (163)
14	Bash Shell	bash (22546)		54	Korn			94	SOSL	
15	BeanShell	beanshell (108)		55	LINC			95	SPARQL	sparql (757)
16	Bourne shell	bourne-shell (123)		56	Lingo	lingo (24)		96	SQL	sql (135777); sql-server (68118); sql-server-2008 (27089)
17	C	c (97364)		57	Lisp	lisp (2657)		97	SystemTap	systemtap (26)
18	C#	c# (468530)		58	LotusScript	lotusscript (257); lotus (299)		98	Tcl	tcl (1639)
19	C++	c++ (204517)		59	Lua	lua (3859)		99	T-SQL	tsql (20534)
20	C++/CLI	cli (1254)		60	Magik			100	TTCN	ttcn (6)
21	CFML	coldfusion (5887); cfml (173)		61	MATLAB	matlab (19466)		101	TypeScript	typescript (967)
22	CLIST	clistctrl (88); clist (10)		62	MUMPS	mumps (24)		102	Uniface	
23	Clojure	clojure (5268)		63	NATURAL			103	VB	
24	COBOL	cobol (423)		64	Objective-C	objective-c (136493)		104	VB.NET	vb.net (47500); vb.net-2010 (897)
25	CoffeeScript	coffeescript (3487)		65	OCaml	ocaml (1622); ocamlbuild (36)		105	VB6	vb6 (5342); vb6-migration (347)
26	C-shell			66	Pascal	pascal (737)		106	VBA	vba (14908); excel-vba (8288); access-vba (1558); word-vba (665); outlook-vba (385); powerpoint-vba (360)
27	Cython	cython (559)		67	PeopleCode			107	VBScript	vbscript (5767)
28	Dart	dart (1063)		68	Perl	perl (27619); perl-module (576); perl6 (61)		108	VC++	visual-c++ (13945); visual-c++-2010 (262); visual-c++-2008 (165); visual-c++-2005 (101); visual-c++-2012 (61); visual-c++-2010-express

								(38); visual-c++-runtime (4)
29	Data Analysis Expressions (DAX)	dax (32); data-analysis (151)	69	PHP	php (396446)	109	WebLogic Scripting Tool	weblogic (2069); weblogic-10.x (501); weblogic11g (248); weblogic12c (95); weblogic9.x (63); weblogic8.x (34); weblogic-integration (7)
30	Data Mining Extensions (DMX)	dmx (18); data-mining (1058)	70	PHP4	php4 (275)	110	X++	x++ (355)
31	DCL	dcl (17)	71	PHP5	php-5.3 (630); php-5.2 (193)	111	XAML	xaml (16402)
32	Delphi	delphi (20976)	72	PL/1		112	XPath	xpath (8601); xpath-2.0 (176); xpath-1.0 (30)
33	ECMAScript	ecmascript-5 (304); ecma262 (104); ecma (56); ecmascript-6 (16); ecmascript-4 (5)	73	PL/SQL	plsql (5371)	113	XQuery	xquery (1421)
34	Elixir	elixir (90)	74	PowerShell	powershell (10273)			
35	Embedded C		75	PROC SQL	proc-sql (97)			
36	Embedded C++		76	Progress 4GL	progress-4gl (183)			
37	Erlang	erlang (3430)	77	Prolog	prolog (2883)			
38	F#	f# (4757)	78	Python	python (192402)			
39	Fortran	fortran (2041)	79	R	r (30644)			
40	FoxPro (VFP)	foxpro (385); vfp (158)	80	REXX	rexx (31)			

Note: Numbers in parenthesis are the number of occurrences of the Stack Overflow tag in the entire Stack Overflow dataset.

Appendix C: Extensions and Further Robustness Checks

This section extends our main analysis and provides additional tests that vary some of the particulars of our empirical design. These extensions and tests are designed to provide ex-post reasoning and thus increase the validity of our empirical strategy, especially the CV-based distinction between career-motivated members and other members while ensuring the robustness of our main results. These analyses also help provide more nuances to our main findings.

Types of Contribution: Questions vs. Answers

We first examine whether the members who we identified as career-motivated optimize their signaling behavior by focusing on contribution activities that promise more signaling gains. Such optimization behavior is to be expected by career-motivated members and demonstrating it will further support our empirical strategy.

As in most Q&A-based online collaboration communities, members in Stack Overflow have two primary ways of contributing: by posting questions and by answering others' questions. Both ways can earn an increase in their reputation score.[20] Thus, we have used the total count of posts, not distinguishing between questions and answers, to measure contributions in the main analyses. However, potential rewards to the member in terms of reputation scores obtainable from a given post vary between questions and answers.[21] While questions get five points each time the question is voted up, answers get ten points for each up-vote plus an additional 15 points if the answer is accepted by the member who originally posted that question. Though answers may typically require greater effort than questions, they also promise a higher expected return in terms of reputation. Because reputation scores are directly connected to members' CVs, we expect that especially career-motivated members are prone to optimize their reputation and are more likely to answer questions than post questions. More precisely, when distinguishing between answers and questions, we expect the positive relationship between career motivation and contribution level, as well as the moderating effects of job-market conditions, to be stronger for answers than for questions. That is in fact what we find.

Table C1. Breakdown by Contribution Type (Questions vs. Answers)

	(1)	(2)	(3)	(4)	(5)	(6)
	Questions			Answers		
(Log) # of days since joining	-0.128**	-0.113**	-0.113**	-0.348**	-0.288**	-0.288**
	(0.003)	(0.004)	(0.004)	(0.004)	(0.007)	(0.007)
(Log) average length of posts	0.409**			0.147**		
	(0.006)			(0.008)		
(Log) career motivation	0.011**			0.031**		
	(0.001)			(0.002)		
(Log) # job vacancies	0.005	0.002	0.001	0.033**	0.034**	0.031**
	(0.006)	(0.006)	(0.006)	(0.010)	(0.011)	(0.011)
(Log) expected salary	-0.011	-0.027	-0.027	0.065	0.043	-0.016
	(0.024)	(0.024)	(0.026)	(0.043)	(0.044)	(0.047)
(Log) career motivation			0.001+			0.003**
× (Log) # job vacancies			(0.001)			(0.001)
(Log) career motivation			-0.000			0.046**
× (Log) expected salary			(0.008)			(0.014)
(Log) career motivation			0.001			0.011**
× (Dummy) transparency			(0.002)			(0.003)
Constant	-1.033**	1.756**	1.758**	1.537**	2.537**	2.560**
	(0.283)	(0.286)	(0.286)	(0.491)	(0.506)	(0.507)
Member FE	No	Yes	Yes	No	Yes	Yes
Language FE	Yes	Yes	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
N	682710	682710	682710	682710	682710	682710
Within R^2	0.077	0.077	0.077	0.086	0.086	0.087

Note: The dependent variable is the (log) number of questions (Models 1-3) or the (log) number of answers (Models 4-6) posted by the member in a given quarter. Robust standard errors, clustered by members, are in parentheses. +, *, ** denotes statistical significance at 10%, 5%, and 1%, respectively.

[20] Members can also make comments or suggest edits to existing posts but the rewards for these activities are either irrelevant to their reputation (comments) or relatively unimportant (edits if accepted).

[21] https://stackoverflow.com/help/whats-reputation (last accessed on August 19, 2019). The scoring system was revised later to assign equal points for both answers and questions, but the extra points for accepted answers were retained.

Table C1 presents the results from replicating Models 1, 2, and 6 of Table 3 (see main analysis) for both questions (Models 1 through 3) and answers (Models 4 through 6). The coefficients indicate that the volume of answers is about three times as sensitive to increases in career motivation as the volume of questions (Models 1 and 4). This suggests that career-motivated members are disproportionately more likely to choose answers over questions. Moreover, all three moderating effects were consistently stronger for answers than questions (Models 3 and 6). In fact, for questions, the coefficient on the interaction term was mostly indistinguishable from zero. The first-order effect of the number of job vacancies remained consistently positive for answers (Models 4 through 6) but insignificant for questions (Models 1 through 3).

Domains of Contribution: Any Programming Language vs. Non-Language vs. Primary Language

Another way to support our identification of career-motivated members and increase the credibility of our empirical strategy is to examine which knowledge domains the contributions of members are directed to. When quantifying member contributions in our main analysis, we counted all posts made by a member regardless of the knowledge domain. But it is possible that career-motivated members concentrate their efforts on specific knowledge domains (i.e., programming languages, or their own primary languages) that may be more representative of their skills to better signal their skills to external recruiters. If so, we would find stronger effects in language domains than in other domains. Hence, we checked whether the effects varied across the contribution domains.

Table C2 shows the results of this analysis. The dependent variable in Models 1 through 3 is the number of posts in any programming language as defined by IT Jobs Watch. Thus, a post is considered a contribution if it belongs to at least one of the 113 programming languages in IT Jobs Watch. The coefficient estimates showed a pattern that was very similar to that in Table 3 (main analysis). Career motivation was positively related to the number of posts in programming languages (Model 1) and all moderators positively interacted with the career motivation-contribution relationship (Model 3). The baseline effect of the *number of job vacancies* was positive, while that of the *expected salary* was insignificant. Thus, our findings from using contributions to all knowledge domains were confirmed on the subset that more strictly defines contribution domains.

Table C2. Breakdown by Contribution Domain (Any Programming Language vs. Non-Programming-Language vs. Primary Programming Language)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Any language			Non-language			Primary language		
(Log) # of days since joining	-0.339** (0.004)	-0.284** (0.006)	-0.284** (0.006)	-0.203** (0.003)	-0.176** (0.006)	-0.176** (0.006)	-0.281** (0.004)	-0.229** (0.006)	-0.229** (0.006)
(Log) average length of posts	0.291** (0.008)			0.183** (0.007)			0.200** (0.007)		
(Log) career motivation	0.027** (0.002)			0.024** (0.002)			0.025** (0.002)		
(Log) # job vacancies	0.041** (0.010)	0.047** (0.010)	0.045** (0.010)	-0.002 (0.008)	-0.014+ (0.009)	-0.017* (0.009)	0.047** (0.009)	0.056** (0.009)	0.054** (0.009)
(Log) expected salary	0.046 (0.040)	0.035 (0.040)	-0.006 (0.043)	0.016 (0.033)	-0.032 (0.034)	-0.071+ (0.037)	0.093* (0.036)	0.065+ (0.037)	0.037 (0.040)
(Log) career motivation × (Log) # job vacancies			0.002* (0.001)			0.003* (0.001)			0.002+ (0.001)
(Log) career motivation × (Log) expected salary			0.032* (0.013)			0.030** (0.011)			0.021+ (0.012)
(Log) career motivation × (Dummy) transparency			0.008** (0.003)			0.006* (0.002)			0.006* (0.002)
Constant	0.549 (0.456)	2.309** (0.469)	2.326** (0.469)	0.896* (0.388)	2.690** (0.405)	2.705** (0.405)	-0.310 (0.420)	1.038* (0.431)	1.050* (0.431)
Member FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Language FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	682710	682710	682710	682710	682710	682710	682710	682710	682710
Within R^2	0.091	0.091	0.091	0.082	0.083	0.084	0.067	0.068	0.068

Note: The dependent variable is the (log) number of posts on any programming language (Models 1-3) or the (log) number of posts on other than programming languages (Models 4-6) or the (log) number of posts on the primary programming language (Models 7-9) by the member in a given quarter. Robust standard errors, clustered by members, are in parentheses. +, *, ** denotes statistical significance at 10%, 5%, and 1%, respectively.

Interestingly, this career-motivated boost in member contributions seems to spill over to non-programming-language-related domains. Models 4 through 6 used the number of posts in non-programming-language domains, computed as the number of total posts minus the number of language-related posts.[22] The results were largely similar to those of programming language domains: career motivation was positively related to the number of posts and all three moderating relationships were positive and significant. The baseline effect of career motivation for non-language posts was smaller than for language posts, though the difference was statistically insignificant. The number of job vacancies and the expected salary had almost no—or even slightly negative—direct relationship with non-language posts. These contrast with their baseline relationships with programming language-related posts. Hence, while confirming our main results on this subset of non-language contributions, we also find suggestive evidence that career-motivated members may allocate more efforts to domains more closely related to their job skills. This becomes more evident when we limited contributions to the posts in the member's primary-language domains (Models 7 through 9). On both job-market parameters, member contributions to their primary-language domains positively responded to more favorable job conditions, in terms of both the direct effect and the interaction effect with career motivation. Together, these results appear consistent with the notion that career-motivated members optimize their effort to maximize returns from private resources to contribute to the community.

Restricted Sample Analysis

To further check the robustness of our results, we tried a few variations of the sample by excluding some members that could have introduced bias in some unknown ways. First, we limited the sample to the members who have posted their CVs (i.e., career-motivated members in our definition). This was to address any potential bias from mismeasurement by treating all members who have not posted their CVs as having any career interest, though some of them may do without posting their CVs.

Limiting the sample to CV members and examining the variations among them will help minimize this bias. The results in Table C3 were consistent with those in the main analysis, with most coefficients supporting the hypothesized relationships, though the results were somewhat weaker for vote-weighted contributions.[23]

Table C3. Contributions of CV Posters

	(1)	(2)	(3)	(4)	(5)	(6)
	Quantity of contributions			Vote-weighted contributions		
(Log) # quarters since first work experience	-0.054** (0.010)	0.079** (0.024)	0.043+ (0.026)	-0.076** (0.012)	0.048 (0.030)	0.033 (0.032)
(Log) # of days since joining	-0.394** (0.006)	-0.345** (0.011)	-0.342** (0.011)	-0.348** (0.007)	-0.326** (0.013)	-0.325** (0.013)
(Log) average length of posts	0.340** (0.012)			0.398** (0.013)		
(Log) career motivation	0.194** (0.007)			0.242** (0.008)		
(Log) # job vacancies	0.042** (0.016)	0.041** (0.016)	0.028 (0.018)	-0.033 (0.021)	-0.030 (0.022)	-0.042+ (0.023)
(Log) expected salary	0.084 (0.064)	0.044 (0.066)	-0.189* (0.091)	0.158+ (0.082)	0.162+ (0.083)	-0.107 (0.116)
(Log) career motivation × (Log) # job vacancies			0.004+ (0.002)			0.003 (0.003)
(Log) career motivation × (Log) expected salary			0.077** (0.023)			0.088** (0.030)
(Log) career motivation × (Dummy) transparency			0.011** (0.005)			0.002 (0.006)
Constant	-0.301 (0.703)	2.759** (0.719)	2.781** (0.717)	-0.277 (0.888)	2.903** (0.914)	2.930** (0.911)
Member FE	No	Yes	Yes	No	Yes	Yes
Language FE	Yes	Yes	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
N	279227	279227	279227	276971	276971	276971
Within R^2	0.097	0.098	0.098	0.145	0.146	0.146

Note: The dependent variable is the (log) number of posts (Models 1-3) or the (log) number of votes (Models 4-6) for the member in a given quarter. Robust standard errors, clustered by members, are in parentheses. +, *, ** denotes statistical significance at 10%, 5%, and 1%, respectively.

[22] Non-language posts are answers or questions that are not associated with any programming-language tags. Examples include operating systems (e.g., "Android" or "iOS"), programming concepts (e.g., "arrays" or "regex"), and development software (e.g., "xcode" or "eclipse").

[23] We also tried further restricting the CV-only sample by imposing different minimum thresholds on the number of words in the CV (e.g., 30 words, 50 words). The results were qualitatively similar though weaker.

Second, we replicated the analysis on a subset of members who joined the community before the initial introduction of the Stack Overflow Jobs (November 3, 2009). These are the members who joined the community before the career service was launched and hence their membership is unlikely to have been influenced by the availability of the career service, which came only later.[24] Table C4 presents the estimates from this subset of early members. The overall pattern of the findings in our main analysis was successfully replicated in the subsample: for both quantity and perceived quality, career motivation was positively related to member contribution and all three moderators strengthened the relationship between career motivation and contribution level. It is noteworthy that the effect of career motivation appears much stronger in this subset, for instance with the magnitude of the quantity effect (0.077; Model 1) more than double that of the entire group (0.032; Model 1 in Table 3, see the main analysis).

Table C4. Contributions of Early Members

	(1)	(2)	(3)	(4)	(5)	(6)
	Quantity of contributions			Vote-weighted contributions		
(Log) # of days since joining	-0.278** (0.010)	-0.174** (0.013)	-0.179** (0.013)	-0.265** (0.013)	-0.182** (0.017)	-0.187** (0.017)
(Log) average length of posts	0.248** (0.014)			0.378** (0.016)		
(Log) career motivation	0.077** (0.003)			0.075** (0.003)		
(Log) # job vacancies	0.039** (0.012)	0.036** (0.013)	0.033* (0.013)	-0.043* (0.017)	-0.046** (0.017)	-0.050** (0.017)
(Log) expected salary	0.102+ (0.053)	0.068 (0.053)	0.019 (0.057)	0.233** (0.069)	0.216** (0.069)	0.167* (0.074)
(Log) career motivation × (Log) # job vacancies			0.003* (0.001)			0.004* (0.002)
(Log) career motivation × (Log) expected salary			0.035* (0.015)			0.035+ (0.020)
(Log) career motivation × (Dummy) transparency			0.013** (0.003)			0.013** (0.004)
Constant	0.334 (0.614)	2.069** (0.627)	2.027** (0.629)	-0.400 (0.778)	2.107** (0.794)	2.063** (0.797)
Member FE	No	Yes	Yes	No	Yes	Yes
Language FE	Yes	Yes	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
N	400774	400774	400774	399028	399028	399028
Within R^2	0.116	0.116	0.117	0.178	0.179	0.179

Note: The dependent variable is the (log) number of posts (Models 1-3) or the (log) number of votes (Models 4-6) for the member in a given quarter. Robust standard errors, clustered by members, are in parentheses. +, *, ** denotes statistical significance at 10%, 5%, and 1%, respectively.

Third, we limited the sample to the members whose location was indicated as one of the EU member countries. The EU ensures free flow of labor across member countries and hence represents one labor market. Though our job-market data from IT Jobs Watch cover only the UK, these jobs are considered equally available to the residents in the EU, albeit subject to their individual costs of relocation. Thus, we repeated our main analysis on the subsample of members from the EU.[25]

The results in Table C5 were qualitatively similar to those in our main analysis, though generally weaker. This may be due to the large heterogeneity across member countries in their socioeconomic infrastructures as well as local job-market conditions, which may deviate more from the trends in the global IT job markets than do larger and more established countries such as the United States and Canada. Nonetheless, it is reassuring to find strong support for the primary effect of career motivation on contributions, while all moderating effects also exhibit correct signs. Overall, these checks on restricted samples largely confirm our main findings.

[24] There are 29,249 unique early members in our sample, which is about 40% of total members.

[25] These countries are Austria, Belgium, Bulgaria, Croatia, Republic of Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, and United Kingdom. Together, the EU members (16,475) represent 22.7% of the total members in our sample.

Table C5. Contributions of EU Members

	(1)	(2)	(3)	(4)	(5)	(6)
	Quantity of contributions			Vote-weighted contributions		
(Log) # of days since joining	-0.339** (0.009)	-0.264** (0.015)	-0.264** (0.015)	-0.306** (0.011)	-0.249** (0.018)	-0.249** (0.018)
(Log) average length of posts	0.363** (0.019)			0.468** (0.022)		
(Log) career motivation	0.038** (0.004)			0.042** (0.005)		
(Log) # job vacancies	0.050* (0.023)	0.054* (0.023)	0.054* (0.024)	-0.033 (0.030)	-0.028 (0.031)	-0.027 (0.031)
(Log) expected salary	0.089 (0.091)	0.067 (0.093)	0.004 (0.098)	0.259* (0.116)	0.269* (0.118)	0.200 (0.125)
(Log) career motivation × (Log) # job vacancies			0.001 (0.003)			0.000 (0.003)
(Log) career motivation × (Log) expected salary			0.053+ (0.030)			0.057 (0.039)
(Log) career motivation × (Dummy) transparency			0.022** (0.006)			0.026** (0.007)
Constant	0.232 (1.014)	2.629* (1.031)	2.606* (1.030)	-0.805 (1.289)	2.071 (1.328)	2.048 (1.327)
Member FE	No	Yes	Yes	No	Yes	Yes
Language FE	Yes	Yes	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
N	168451	168451	168451	167680	167680	167680
Within R^2	0.097	0.097	0.098	0.150	0.150	0.151

Note: The dependent variable is the (log) number of posts (Models 1-3) or the (log) number of votes (Models 4-6) for the member in a given quarter. Robust standard errors, clustered by members, are in parentheses. +, *, ** denotes statistical significance at 10%, 5%, and 1%, respectively.

Nonlinearity of Baseline Effect

So far, we have used CV length as a continuous measure of the degree of career motivation. A simpler way of capturing the effect of career motivation would be to use a dichotomized variable that distinguishes between members who have posted their CVs and those who have not. However, that would amount to assuming that all members who have posted their CV are *equally* motivated by a career, which might be an oversimplification. Members who simply explored the CV feature would be considered career-motivated although they might have actually been driven by curiosity. In fact, the raw correlation between the number of posts and the CV dummy is slightly negative (-0.038), opposite to our results based on the continuous measure. A closer examination of CV length reveals that the distribution of CV length is roughly bimodal, with one large cluster at a very low level and the other around a relatively high level. The shortest CVs consist of only a few words, hardly useful to recruiters. In contrast, CVs in the other cluster are long enough (700-1100 words) to present information useful for signaling skills. This suggests that the degree of career motivation represented by CV length may be discontinuous among the CV members, with some members very serious about pursuing a career, while others are not so serious or less dedicated than even the no-CV members. If so, our estimation predicated on the continuity of the degree of career motivation among the CV member group may underestimate the true effect of career-motivated member contributions.

Table C6. Nonlinearity of Baseline Effects (Random-Effects Models)

	(1) Quantity	(2) Vote-weighted	(3) # Questions	(4) # Answers	(5) # Posts in Prog. Lang.	(6) # Posts by Early Members	(7) # Posts by EU Members
(Log) # of days since joining	-0.376** (0.004)	-0.337** (0.005)	-0.132** (0.003)	-0.370** (0.004)	-0.355** (0.004)	-0.270** (0.010)	-0.356** (0.009)
(Log) average length of posts	0.335** (0.008)	0.419** (0.009)	0.405** (0.006)	0.129** (0.008)	0.278** (0.007)	0.232** (0.013)	0.344** (0.018)
(Dummy) low CM	-0.289** (0.007)	-0.371** (0.008)	-0.047** (0.005)	-0.320** (0.007)	-0.227** (0.006)	-0.247** (0.011)	-0.300** (0.016)
(Dummy) high CM	0.342** (0.014)	0.383** (0.016)	0.111** (0.009)	0.338** (0.014)	0.282** (0.013)	0.502** (0.017)	0.354** (0.027)
(Log) # job vacancies	0.025* (0.011)	-0.055** (0.014)	0.005 (0.006)	0.031** (0.010)	0.040** (0.010)	0.038** (0.012)	0.049* (0.023)
(Log) expected salary	0.069	0.155**	-0.009	0.074+	0.053	0.100+	0.101

	(0.044)	(0.056)	(0.024)	(0.043)	(0.040)	(0.053)	(0.091)
Constant	0.795 (0.506)	0.815 (0.633)	-0.995** (0.283)	1.737** (0.489)	0.701 (0.454)	0.554 (0.612)	0.352 (1.009)
Member FE	No	No	No	No	No	No	No
Language FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	682710	678139	682710	682710	682710	400774	168451
Within R^2	0.102	0.156	0.077	0.085	0.09	0.116	0.096

Table C7. Nonlinearity of Baseline Effects (Fixed-Effects Models)

	(1) **Quantity**	**(2)** **Vote-weighted**	**(3)** **# Questions**	**(4)** **# Answers**	**(5)** **# Posts in Prog. Lang.**	**(6)** **# Posts by Early Members**	**(7)** **# Posts by EU Members**
(Log) # of days since joining	-0.283** (0.007)	-0.272** (0.009)	-0.113** (0.004)	-0.288** (0.007)	-0.283** (0.006)	-0.179** (0.013)	-0.263** (0.015)
(Log) # job vacancies	0.023* (0.011)	-0.058** (0.015)	0.001 (0.006)	0.032** (0.011)	0.045** (0.010)	0.032* (0.013)	0.049* (0.024)
(Log) Expected salary	0.008 (0.050)	0.141* (0.064)	-0.017 (0.027)	0.011 (0.049)	0.022 (0.045)	0.057 (0.060)	0.049 (0.100)
(Dummy) Low CM × (Log) # job vacancies	0.006 (0.005)	0.011+ (0.006)	0.004 (0.003)	0.004 (0.005)	0.003 (0.004)	0.008 (0.006)	0.014 (0.011)
(Dummy) high CM × (Log) # job vacancies	0.020* (0.009)	0.025* (0.011)	0.007 (0.005)	0.022* (0.009)	0.015+ (0.008)	0.018* (0.009)	0.003 (0.018)
(Dummy) low CM × (Log) expected salary	-0.041 (0.054)	-0.063 (0.069)	-0.036 (0.031)	-0.000 (0.052)	-0.034 (0.047)	-0.064 (0.065)	-0.103 (0.116)
(Dummy) high CM × (Log) expected salary	0.297** (0.098)	0.325* (0.127)	0.010 (0.052)	0.319** (0.097)	0.238** (0.087)	0.237* (0.100)	0.391* (0.197)
(Dummy) low CM × (Dummy) transparency	0.008 (0.011)	0.041** (0.014)	0.009 (0.006)	0.026* (0.011)	0.008 (0.010)	0.016 (0.012)	-0.011 (0.023)
(Dummy) high CM × (Dummy) transparency	0.089** (0.019)	0.083** (0.024)	-0.000 (0.010)	0.077** (0.019)	0.064** (0.017)	0.094** (0.020)	0.167** (0.038)
Constant	3.068** (0.523)	3.253** (0.656)	1.755** (0.286)	2.560** (0.507)	2.326** (0.469)	1.989** (0.631)	2.659** (1.021)
Member FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Language FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	682710	678139	682710	682710	682710	400774	168451
Within R^2	0.103	0.157	0.077	0.087	0.091	0.117	0.098

To examine this possibility, we divided the CV member group into a "low career motivation" group (53 words or fewer in the CV) and a "high career motivation" group (more than 53 words), where the 53-words count roughly equally divides the bimodal distribution. We then reestimated all models with each career motivation group dummy representing one of the two groups. No career motivation (having no CV) is the base category here. Tables C6 and C7 present the results from these analyses. The random-effects models of Table C6 testing the effects of the career motivation dummies for the high-motivation members exhibited a strongly positive coefficient across all models. The models differ in the measure of member contributions or the sampling. The magnitude of the effects was also sizable: compared to the no-motivation group, the high-motivation group's quarterly contribution volume was 30-70% higher. Even in the question category, on which the career-driven boost in contribution volume was relatively muted (Table C1), the high-motivation group contributed 11% more than the no-motivation group. This result further increases confidence in our findings of career-motivated contribution behavior. In contrast, for the low-motivation group, the coefficient was significantly negative in all models. This indicates that the members who have posted grossly incomplete CVs contribute even less than those who did not post their CVs at all, not to mention those who posted more complete CVs. These short CVs typically contain no more than the member's name and location. These members may have simply dabbled with the new feature without any serious intent of using it for a career-related purpose. Thus, any effect we find on the role of career motivation on member contributions is likely to be driven by the high-motivation group of members (which is also consistent with the results on the CV-only sample in Table C3). These are the members who likely have spent time and effort to fill out their resumes as completely as possible and perhaps even updated to make them current. Results from the fixed-effects model (Table C7) also show that the moderating effects of job-market conditions were entirely driven by the high-motivation members; for the low-motivation members, none of these moderators appeared to influence the association between career motivation and contribution level. This analysis reveals stark heterogeneity among members who posted their CVs. Hence, our estimates based on the continuous measure that essentially treats the members with very short CVs as also career-motivated, albeit to a lesser extent, are likely to be conservative.

Appendix D: Coarsened-Exact Matching Method

To address possible endogeneity, we created a matched sample between CV members and no-CV members using the coarsened-exact matching (CEM) method. CEM is a method that improves causal inference by creating a matched sample that imposes comparable equality between the treated group and the control group along multiple dimensions of interest simultaneously [35]. While exact matching will be ideal, it is practically impossible to exactly match on a continuous variable. CEM first coarsens the continuous variable into multiple bins and then enforces equality between groups within each bin. Because different numbers of controls may be matched with each treated observation across bins, CEM generates the weights that can be used in regression analysis to estimate the average treatment effect on the treated. For the implementation, we used the CEM algorithm in Stata 15 [12].

In our context, posting a CV (dummy) is the treatment. Because CVs are time-invariant in our data, our matching process used cross-sectional data of members by collapsing the panel data. We employed three member-level variables for matching members between two groups: (1) primary programming language, (2) log number of total posts, and (3) log number of total votes received. We imposed an exact matching for the programming language, which is a categorical variable while coarsening the numbers of posts and votes using Scott's rule for binning [56]. In total, we obtained 6,018 matched strata, where 62,358 members (86%) belonged to one of these strata. The comparison between the pre-CEM balance (Table A4-1) and the post-CEM balance (Table A4-2) suggests successful matching, as following the CEM, the imbalances between CV members (treated group) and no-CV members (control group) in the matching variables all but disappeared.

Table A4-1. Imbalances before CEM Matching

Variable	L1 Distance	Mean	Min	25%	50%	75%	Max
Primary programming language	0.037	0.554	0.000	0.000	0.000	0.000	0.000
(Log) # Posts	0.139	-0.453	0.000	-0.496	-0.415	-0.366	-0.652
(Log) # Votes	0.159	-0.647	0.000	-0.773	-0.627	-0.546	-1.152

Table A4-2. Imbalances after CEM Matching

Variable	L1 Distance	Mean	Min	25%	50%	75%	Max
Primary programming language	0.000	0.000	0.000	0.000	0.000	0.000	0.000
(Log) # Posts	0.010	0.000	0.000	0.000	0.000	0.000	-0.027
(Log) # Votes	0.008	0.000	0.000	0.000	0.000	0.000	-0.018

About the Authors

Jeongsik "Jay" Lee is an associate professor of management at the LeBow College of Business, Drexel University. Prior to joining Drexel, he was on the faculty of the Scheller College of Business, Georgia Institute of Technology. He received a PhD in management from the University of California, Los Angeles. He also holds a Master of Business Administration degree from Duke University and a Bachelor of Arts degree in economics from Seoul National University, Korea. Before pivoting to academia, Dr. Lee worked at the Bank of Korea as an economist, analyst, and manager. Dr. Lee's research focuses on issues involving technological innovation, knowledge management, and social networks. His work has appeared in leading management journals such as *Academy of Management Journal*, *Management Science*, *MIS Quarterly*, *Organization Science*, *Research Policy*, and *Strategic Management Journal*.

Hyunwoo Park is an associate professor in the Graduate School of Data Science at Seoul National University. Before joining SNU, he was an assistant professor in operations and business analytics at the Fisher College of Business and a core faculty member in the Translational Data Analytics Institute (TDAI) at The Ohio State University. Prior to OSU, he was a postdoctoral fellow at the Tennenbaum Institute at Georgia Tech. He holds a PhD in industrial engineering from Georgia Tech, a Master of Information Management and Systems degree from UC Berkeley, and a BS in electrical engineering from Seoul National University. His research interests include business and data analytics with an emphasis on visualization, supply network management from the network perspective, and technology and innovation management in the presence of digital platforms. His research has appeared in leading journals including *Academy of Management Review*, *Decision Sciences*, *Decision Support Systems*, *IEEE Transactions on Engineering Management*, *IEEE Transactions on Visualization and Computer Graphics*, *Journal of Operations Management*, and *Research Policy*. His work has been recognized by several awards and nominations by the Academy of Management and INFORMS.

Michael Zaggl is an associate professor in the School of Business and Social Sciences at Aarhus University, Denmark. Before joining Aarhus University, he led a research group at the Technical University of Munich, TUM School of Management, where he also received a Habilitation degree. Prior to TUM, he received a PhD in computational economics from the Technical University of Hamburg and an MSc and BSc in computer science and management from the University of Koblenz. His research includes distributed and digital innovation, crowdsourcing, technologically enabled search, and data-driven decision-making. His research has been published in leading journals, e.g., *Academy of Management Discoveries*, *IEEE Transactions on Engineering Management*, and *Research Policy*.