Open Source Software and Innovation in the Software Development Industry – Draft

Nathalie Rosalind Tate

Oakland University

Abstract

This paper examines the ways in which open source software (OSS) is better for innovation in the software development industry than proprietary software. An examination of prior research and two surveys of software developers and computer science students clearly indicate that the production and availability of OSS benefits the software developer.

Keywords: open source, proprietary, software, innovation

Open Source Software and Innovation in the Software Development Industry – Draft

Open source software is software where the source code is freely available. Although the exact restrictions vary between open source licenses, most allow for users to edit and redistribute the software. One question raised by the rise of open source software (OSS) is how it compares to proprietary software with regard to innovation in the software industry, where "innovation" is defined as a holistic metric that considers the amount of new software and code commits, bug reports, and bug fixes to existing software, as well as the perceived quality and usability of said software.

The main divide between proprietary software and OSS, other than the availability of the source code, is the price: Proprietary software frequently must be purchased, while OSS is usually provided for free. This price difference makes it easier to monetize proprietary software and is a large factor in the popularity of proprietary software.

Literature Review

Atal and Shankar (2014) compares how price competition and differences in software quality vary between two markets: one in which two proprietary software firms are competing and one in which a proprietary software firm competes against an open source software (OSS) firm. The authors utilized a statistical model to compare the two types of markets. The models compare software quality and price competition. Although OSS is often free, the authors factored in an "effort cost" when determining the price competition. The findings of the model were that a proprietary firm competing against an OSS firm may have either higher or lower profits than a proprietary firm competing against another proprietary firm, depending on the price competition from the OSS firm. The price competition depends on the effort cost of the OSS, the degree to which consumers value quality, and the degree to which the consumers of the OSS contribute to the production of the OSS. The proprietary firm had the highest profits when the price competition from the OSS firm was low, but the consumers were better-off with lower price competition. The

authors also found that OSS had low value for users who valued quality in their software, but that it may have high value for users who did not value quality (Atal & Shankar, 2014).

Howison and Bullard (2016) discuss the visibility of software used and mentioned in biology literature. The authors compiled a "balanced and representative sample" of biology literature and analyzed the content by looking for mentions of software. Once found, the mentions were "coded" according to a coding scheme defined by the authors. The coding scheme contained information about the mentioned software: if the mention was cited, if the mentioned software was easily identifiable and findable, if the source code was available, and if the author of the software granted permission for others to edit it. The authors found that they had access to 79% of the mentioned software, 47% was available for free, 32% allowed access to the source code, and 20% allowed users to modify the code. The authors concluded that the manner in which software was mentioned and cited in scientific literature are "varied and appear relatively ad hoc." The authors end with a call for more research (Howison & Bullard, 2016).

Sacks (2015) utilizes a statistical model to examine competition between a firm that produces proprietary software and a community that produces OSS. The author's findings are that the firm will cater to the less technologically savvy and the community will cater to those who are more technologically capable. In the presence of competition from the community, the firm targets "even less technologically savvy individuals". Furthermore, the author notes that the "real price" of software has two factors: the purchase price and the opportunity cost of learning how to use it. Additionally, the author comments on how proprietary firms can benefit from "R&D spillover" from permissively licensed OSS projects (Sacks, 2015).

Widenius and Nyman (2014) discuss the "business of open source software". The authors discuss possible business models that allow an individual or corporation to profit from OSS. These business models include: "Support contracts and services", where one only supplies support for the software; "Open core", where the core software is OSS but the

peripheral software and extensions are closed source; “Business source”, where the software is maintained under two licenses with a time delay; “Dual licensing”, where the software is maintained under a proprietary license and a strong copy-left license, so customers who cannot or chose not to comply with the terms of the copy-left license must buy the software; and “Software as a service”, where the application software is OSS, but the server it communicates with is closed source (Widenius & Nyman, 2014).

Xing (2014) examines the effects of both free and commercial OSS on proprietary software. The author used a statistical model to compare two different hypothetical markets: one where proprietary software only competes with free OSS and another where proprietary software competes against both free and commercial OSS. The model considered differences in usability between OSS and proprietary software, as well differences in functionality and user technical ability. The authors findings were that the introduction of commercial OSS to the market can decrease the price of proprietary software and increase social welfare without necessarily decreasing the market share of the proprietary software. A market in which proprietary software only competes against free OSS, however, may lead to higher prices and lower social welfare (Xing, 2014).

Methods

Two surveys were created using SurveyMonkey.com: one was posted on the popular programming subreddit /r/coding on reddit.com; the other was distributed via Slack to the employees of Xby2, a software development company in Farmington Hills, MI. Slack is an instant-messaging platform frequently used by software developers. The surveys contain the same questions, but have been kept separate to allow for comparing the data between the two sets of respondents.

Xby2 employs between 51 and 200 people (“Xby2,” n.d.), and the survey is expected to reach approximately 40 people. The gender makeup of Reddit is 53% male and 47% female. Additionally, 54% of users are American and 46% are international. (“Audience

and Demographics," n.d.). /r/coding was chosen over the more popular /r/programming specifically because it is less popular. As of 2016-10-17, /r/programming has 699,453 subscribers and /r/coding has 75,898. /r/coding has more strictly enforced rules and a smaller community, which could lead to higher quality sample data.

The survey contains six questions. The first question ensures that the respondent is in the target demographic and the remaining five questions ask the respondent about the Open Source and Proprietary projects they work on, the licenses they use (if Open Source), and the types of software they use. The licenses and types of software used were chosen to be a representation of some of the most commonly used, with an option for "Other" to account for omissions.

Results

Discussion

Atal and Shankar (2014) and Sacks (2015) both compare a market where two proprietary firms compete to one in which a proprietary firm competes with a not-for-profit OSS firm. An important difference is that while Atal and Shankar (2014) consider an "effort cost" to producing OSS (Atal & Shankar, 2014, p. 334), Sacks (2015) instead considers a "learning cost" to using it (Sacks, 2015, p. 271). An effort cost is the cost for the developers to produce the software. The reasoning here is that as the OSS is free, the developers are not paid and therefore incur a net loss in producing the software. While this is certainly true in some cases, it fails to account for paid OSS developers or OSS developers who genuinely enjoy their work. My primary research supports the idea of an effort cost, with 58% of Reddit respondents reporting that they rarely or never contribute to OSS, while 55% reported contributing to proprietary projects more than once a week. The learning cost on the other hand is the cost for the user of the OSS to learn how to use it. My primary research does not support the idea of a learning cost being prohibitive. In fact, in the Reddit survey, OSS beat proprietary software for number of users in every category

except for Office Suites. This discrepancy is likely due to the fact that everyone polled in the Reddit survey was a either software developer or computer science student, and these demographics are likely to have higher technical knowledge and lower learning costs. A poll of the general public would likely have the expected results. Sacks (2015) accounts for these potential discrepancies in their model by acknowledging that any two users may have drastically different learning costs. Sacks (2015) also considers R&D spillover from OSS communities into proprietary firms, which is an advantage of their model.

That both Atal and Shankar (2014) and Sacks (2015) both only consider free OSS is significantly disadvantageous for both models. Xing (2014) on the other hand, considers the effects of both free and for-profit OSS on proprietary software. Like Sacks (2015), Xing (2014) also considers a learning cost for the users of the OSS, but the model neither considers an effort cost nor R&D spillover.

R&D spillover, as detailed by Sacks (2015), relies on permissively licensed OSS. Permissive software licenses like the MIT, Apache, BSD, and WTFPL licenses allow OSS to be used in proprietary projects. This is in contrast to copyleft licenses like the GNU GPL, which state that all derivative works must be licensed under a similar license. 58% of surveyed Reddit users reported using one or more of the mentioned permissive licenses at least once, compared to 44% who reported using the GPL. This trend towards permissive licencing could positively impact the amount of R&D spillover that occurs, thereby increasing the amount of new software being developed by companies and having a positive effect on innovation.

Another factor that affects R&D spillover other than the licensing of the OSS is the existence of the OSS itself. Howison and Bullard (2016) explain how different factors affecting the visibility of code in scientific literature, including the availability of the source code, affect scientific research. From Howison and Bullard (2016)'s findings that the availability of source code is beneficial to scientific research and Sacks (2015)'s findings that permissively licensed OSS is beneficial to proprietary software, a logical conclusion

would be that the existence of more OSS would be beneficial to proprietary software companies via the process of R&D spillover.

While important for individuals and not-for-profit OSS communities, choice of software licenses is much more important for for-profit OSS companies. As explained in Widenius and Nyman (2014), many of the more effective ways to monetize OSS require the owner to be very deliberate in their choice of license. For example, "dual licensing" depends on the open source version of the code being released under the GPL or a similar copyleft license. The success of commercial OSS is important beyond the success or failure of any single company: as Xing (2014) found, commercial OSS is often of higher quality than free OSS, and its introduction to the market lowers prices and increases social welfare.

References

Atal, V. & Shankar, K. (2014). Open source software: Competition with a public good. *Atlantic Economic Journal*, *42*(3), 333–345. doi:10.1007/s11293-014-9426-2

Audience and demographics. (n.d.). *Reddit Help*. Retrieved from http://reddit.zendesk.com

Howison, J. & Bullard, J. (2016). Software in the scientific literature: Problems with seeing, finding, and using software mentioned in the biology literature. *Journal of the Association for Information Science and Technology*, *67*(9), 2137–2155. doi:10.1002/asi.23538

Sacks, M. (2015). Competition between open source and proprietary software: Strategies for survival. *Journal of Management Information Systems*, *32*(3), 268. doi:10.1080/07421222.2015.1099391

Widenius, M. & Nyman, L. (2014). The business of open source software: A primer. *Technology Innovation Management Review*, *4*(1), 4. Retrieved from http://www.proquest.com

Xby2. (n.d.). *Glassdoor*. Retrieved from http://glassdoor.com

Xing, M. (2014). The impact of commercial open source software on proprietary software producers and social welfare. *Journal of Industrial Engineering and Management*, *7*(5), 1183–1196. doi:10.3926/jiem.1260