DISCUSSION PAPER SERIES

No. 8434

A BRIEF SURVEY OF THE ECONOMICS OF OPEN SOURCE SOFTWARE

Chaim Fershtman and Neil Gandal

INDUSTRIAL ORGANIZATION

Centre for Economic Policy Research

www.cepr.org

Available online at: **www.cepr.org/pubs/dps/DP8434.asp**

ISSN 0265-8003

A BRIEF SURVEY OF THE ECONOMICS OF OPEN SOURCE SOFTWARE

Chaim Fershtman, Tel Aviv University and CEPR
Neil Gandal, Tel Aviv University and CEPR

Discussion Paper No. 8434
June 2011

Centre for Economic Policy Research
77 Bastwick Street, London EC1V 3PZ, UK
Tel: (44 20) 7183 8801, Fax: (44 20) 7183 8820
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 8434

June 2011

ABSTRACT

A Brief Survey of the Economics of Open Source Software*

The open source model is a form of software development in which the source code is made available, free of charge, to all interested parties; further users have the right to modify and extend the program. Open source software (OSS) methods rely on developers who reveal the source code under an open source license. Under certain types of open source licenses, any further development using the source code must also be publicly disclosed. In this brief survey, we will focus on several key aspects of open source software.

JEL Classification: L17
Keywords: digital content, intrinsic motivation, licenses, open source software and R&D

Chaim Fershtman
Eitan Berglas School of Economics
Tel-Aviv University
Tel Aviv 69978
ISRAEL

Email: fersht@post.tau.ac.il

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=104930

Neil Gandal
Department of Public Policy
Tel Aviv University
Tel Aviv 46766
ISRAEL

Email: gandal@post.tau.ac.il

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=144974

* We are especially grateful to Shane Greenstein for many comments and suggestions that significantly improved the manuscript. We are also grateful to Jacques Lawarree, and Nick Tsilas for very helpful comments and suggestions. An academic research grant from Microsoft is gratefully acknowledged. Any opinions expressed are those of the authors.

Submitted 01 June 2011

1. Introduction:

The open source model is a form of software development in which the source code is made available, free of charge, to all interested parties; further users have the right to modify and extend the program. Open source software (OSS) methods rely on developers who reveal the source code under an open source license. Under certain types of open source licenses, any further development using the source code must also be publicly disclosed.

The open source model has become quite popular and is often referred to as a movement with an ideology and enthusiastic supporters. See for example Stallman (1999) and Raymond (2000). At the core of this process are two interesting phenomena: Unpaid volunteers do a non-trivial portion of the development of open source programs and, unlike commercial software, open source software is not sold or licensed for a fee.

Having unpaid volunteers develop 'free' software is a puzzling phenomenon for economists.[3] What are the incentives that drive contributors to invest time and effort in developing these open source programs, which are not sold or licensed for a fee? Intrinsic motivation may provide a partial explanation and suggests an analogy between academia and the open source movement. While publication plays an important role in academia, the analogy in the open source software (OSS) world is being included in the "list of contributors" of different projects. Being listed as a contributor may enhance the reputation of a programmer and can be instrumental in the job market. Additional incentives to develop open source software come from 'self-use' benefits and the enhancement of other (potentially proprietary) products in the market.

In this brief survey, we will focus on several key aspects of open source software. Much of the empirical work we review in this survey paper comes from high-quality data on open source software projects which are publicly available. Since most open source development takes place in the public domain,[4] data on many

[3] Boldrin, M., and D. Levine (2009) argue that from a historical perspective, the "open source" model of development is the norm for many industries. In this entry, we will focus on the open source phenomenon in software. See section 7 for extensions of open source methodology to other applications.

[4] When we refer to the public domain, we mean publicly available 'via the Internet.'

aspects of open source development are often available at various forges or platforms. These forges typically host many independent software projects. SourceForge, the largest forge, had more than 240,000 projects and 2.6 million registered users as of August 2010. Analyzing the open source data available at SourceForge.net has already provided insight on worker motivation, the tradeoffs between intrinsic and monetary motivation, and the effect of the form of licensing on the contributions of developers.[5]

In section 2, we examine motivation of programmers, while section 3 examines the types of licensing employed in open source projects. Sections 4 and 5 consider changes in the open source model. In section 4, we examine firm participation in the open source process, while in section 5 we review some changes in the institutional structure of open source.

Open source development leads to very different incentives for R&D development than the traditional proprietary development model. See Maurer and Scotchmer (2006) for a detailed analysis. Hence, examining open source successes and failures may shed some light on the R&D process itself. We briefly examine this issue in section 6. In section 7, we briefly discuss the extensions of open source software model to digital content.

Finally, this is a short review; hence we focus on the topics we consider to be most important. Several books provide detailed reviews of open source software. See Dibona et al. (1999, 2006), and Lerner and Schankerman (2010.) Excellent early survey articles include Lerner and Tirole (2005a) and von Krogh and von Hippel (2006).

2. Motivation of Programmers

2.1 Theoretical Research of the Motivation of Programmers

Early research on the open source phenomenon was primarily theoretical and focused on the motivation of unpaid programmers to work on open source projects. Several explanations regarding motivation have been offered in the literature: Lerner and Tirole (2002) argue that developers of open source programs acquire a reputation,

[5] See Lerner and Tirole (2005b) and Fershtman and Gandal (2007), which are discussed in section 3.

which is eventually rewarded in the job market, while Harhoff, Henkel and von Hippel (2003) argue that end users of open source benefit by sharing their innovations. Ghosh et al (2002) argue that open source development is more like a hobby than a (paying) job. Johnson (2002) develops a model of open source software as voluntary provision of a public good – but for such a model one needs to assume that the primary motivation of developers is the "consumption" or the use of the final program.[6]

2.2 Empirical Research of the Motivation of Programmers

Using survey methods, Hars and Ou (2001) and Hertel, Niedner, and Herrmann (2003) find that peer recognition and identification with the goals of the project are the main motivations for developers who contribute to open source software projects. In particular, Hars and Ou's (2001) survey conducted among OSS programmers revealed that peer recognition was an important motivating factor for 43 percent of the respondents, while community identification was a key factor for 28 percent of the respondents. Similarly, Hertel, Niedner, and Herrmann (2003) survey of 141 contributors to Linux kernel project and find that prime motivating factor is "identification with Linux kernel."

Hann, Roberts, Slaughter, and Fielding (2004) empirically examine the Apache HTTP Server Project and find that contributions are not correlated with higher wages, but a higher ranking within the Apache Project is indeed positively correlated with higher wages. Using a Web-based survey Lakhani and Wolf (2005) found that intrinsic motivations help induce developers to contribute to OSS. Chakravarty, Haruvy, and Wu (2007) found that the motivation of OSS programmers depends both on private motivations (like future monetary payoffs or ego) and social motivations (like altruism).

3. Licensing of Open Source Software

Like other products based on intellectual property, the intellectual property in software is typically "licensed" for use, not sold outright. This is the case regardless of whether the software is proprietary or open-source. Even though open source

[6] Johnson (2006) presents a model in which the OSS organization structure is superior to that of proprietary development as it minimizes transaction costs and avoids agency problems.

software is distributed freely without payment, the programs are distributed under licensing agreements. There are several different types of open source licenses. The main difference is the degree of restrictions they entail.

Reciprocal (or viral) licenses require that modifications to the program also be licensed under the same license as the original work. Examples of reciprocal licenses are GNU General Public License (GPL) and the GNU Lesser General Public License (LGPL). The most popular open source license is the GPL. If a software program is distributed under a GPL, the source code must be made available to users. Further, programs that incorporate code from a software project employing a GPL also must insure that the source code is available. The GPL is, hence, a very restrictive license and it is difficult to develop commercial products under a GPL license.[7]

More permissive (non-viral) licenses enable redistribution under a small set of rules. Under these licenses, the software can be modified without making the new source code available publicly as long as the proper attribution is given. Examples of such licenses include the Berkeley Software Development (BSD) license, the Apache License and the Mozilla public License. Commercial products can be developed using software licensed under a BSD-type license as long as credit for the underlying code is given to the copyright holder(s).

Many of the open source programs employ restrictive licenses that would seem to hinder commercial development, since these licenses require that all 'future' software using the relevant code must also be in the public domain.

Several papers in the literature have empirically examined the effect of different licenses. Bonaccorsi and Rossi (2002) surveyed Italian firms that use open source software and found that, on average, firms that employ software with restrictive licenses supply fewer proprietary products than firms that employ software with less restrictive licenses.

The remaining papers we survey in this section come from the very detailed data that are publicly available at SourceForge. Project-level data include the "names" of contributors, their role in the project, who contributed each part of the code, when the development took place, the stage of development, communications among

[7] The LGPL is also quite restrictive, but less restrictive than the GPL restrictive license.

project members, how bugs were fixed, how many times the projects was downloaded the intended audience, type of license, operating system, etc.

Lerner and Tirole (2005b) examine the choice of licenses using the database of open source projects from the SourceForge web site. They find that open source projects that run on commercial operating systems and projects that are designed for developers tend to use less restrictive licenses, while projects that are targeted for end users tend to use more restrictive licenses.

Fershtman and Gandal (2007) find that output per developer is much higher in OSS projects with less restrictive licenses. This is striking since the type of license does not "technically" affect the writing of the code. This result is consistent with the hypothesis that the main motivation of programmers to contribute to restrictive OSS projects is to be included in the "list of contributors:" programmers have a strong motivation to contribute until the threshold level, and weak motivation to contribute above that level. Comino, Manenti and Parisi (2007) find that the more restrictive the license, the lower the probability that the project will reach an advanced development stage.

4. Changes in the Open Source Model – Firm Participation

4.1 Increased Firm Participation in Open Source Projects

The degree of reliance on unpaid programmers has changed over time. More of the work on open source projects is done by contributors who work for firms. Employing a sample of 100 open source projects hosted at Sourceforge.com., Lerner, Parag and Tirole (2006) find that the share of corporate contributors is higher in larger open source projects, where large means more lines of code.

4.2 Open Source & Proprietary Software in Same Market

Several open source products have had great success. Indeed, in most software markets, open source and proprietary products compete side by side. In many of these

markets, open source products have a non-trivial market share as the following examples show:[8]

- Web Browsers: According to W3Counter, in September 2009, Firefox (which is an open source software program) had 32% of the web browser market.[9]
- Web Servers: Apache (which is an open source software program) has been the dominant firm in this market for many years. As of September 2009, Apache served approximately 55% of all websites.[10]
- Server Operating System Market: According to IDC (2008), as cited by Llanes and De Elejalde (2009), Linux had 13.7% of server operating system market.[11]
- According to Trefis,[12] 'MySQL,' an open source database software program, had approximately a 20% market share in database installations worldwide in 2010.

4.3 Towards Mixed-Source Strategies

A key change over time in the open source model is that many proprietary firms now initiate open source projects themselves, in addition to supplying programmers. Indeed, many proprietary firms now use a mixed-source model, that is, a model in which some of their products are proprietary and are distributed under traditional licenses, while some of their products are open source and distributed under an open source license. Such a mixed-source strategy enables firms to benefit from the advantages of both open source and proprietary development. One key advantage to open source software development is that because the code is developed in the public domain, problems (bugs) will be found and solved quickly.

In a huge survey of more than 2300 companies in fifteen countries, Lerner and Schankerman (2010) found that more than 25% of all firms surveyed develop both open source and proprietary software programs. Using data on 73 Finnish software

[8] Recent theoretical work examines this phenomenon as well. See for example Casadesus-Masanell and Ghemawat (2006), Economides and Katsamakas (2006), Athey and Ellison (2009), and Llanes and De Elejalde (2009).

[9] The market data are from W3Counter and are from October 2009. See http://www.w3counter.com/globalstats.php, accessed November 3, 2009.

[10] See http://en.wikipedia.org/wiki/Apache_HTTP_Server, accessed November 3, 2009.

[11] According to W3Counter, in the overall operating system market, Linux held a slightly larger than two percent share. See http://www.w3counter.com/globalstats.php,

[12] See Trefis analysis at https://www.trefis.com/company?article=12891#, accessed January 31, 2011.

companies, Koski (2005) empirically examined which factors affect whether the firm releases its product using an OSS or proprietary licenses. She found that the more the firm is service oriented the more it will be likely to offer products using OSS licenses.

5. Institutional Changes in the Open Source Model[13]

Rules for participation and governance in open source software projects have changed as well over time. Initially, open source projects were rather informal organizational processes. While some open source projects still allow unrestricted participation, many do not. In addition to rules regarding participation, open source projects typically have rules for deciding versions, and rules about reuse.

The institutional setting in which open source takes place has also evolved over time. Sourceforge, which we discussed above, is clearly not the only setting in which open source occurs. Indeed, Sourceforge is an ideal platform when an open source project lacks an institutional home. But, there are many important cases in which open source projects are hosted within an institutional setting. Linux operates within a consortium supported by many firms – and senior personnel receive salaries from the organization. In other cases, firms sponsor open source projects – Webkit, which received financing from Apple, is an example.[14]

Open source has also become a part of standard development by standard setting organizations (SSOs.) The Internet Engineering Task Force (IETF) is essentially both an open source organization and an SSO.[15]

6. Open Source Software and Incentives for R&D[16]

Incentives for engaging in R&D are quite different under open source software than under traditional proprietary software. Under the latter development method, products are often protected by patents and copyrights, which do not typically require disclosure of the source code. Hence, intellectual property laws provide protection

[13] This section draws heavily from Greenstein (2011) and comments and suggestions made by Greenstein.
[14] See West and O'Mahoney (2008) for work in this area.
[15] See Bradner and Scott (1999).
[16] This section draws from Maurer and Scotchmer (2006).

against imitation. Since open source software is typically put into the public domain, open source software would not provide innovation incentives when the goal is to prevent imitation.

However, as Bessen and Maskin (2006) note, imitation of a discovery can be desirable in a world of sequential/cumulative and complementary innovation because it helps the imitator develop further inventions. Since a non-trivial amount software innovation is either sequential/cumulative or complementary (or both,) it suggests that the open source development method may be socially preferable. Interestingly, Maurer and Scotchmer (2006) argue that open source development can also be privately preferable to traditional intellectual property protection when innovation is either sequential/cumulative or complementary.

Open source development also has implications for the cost of R&D. Open source development can be thought of as 'pooled' R&D, which typically implies cost savings - see West and Gallagher (2006.) Firms share code to test software, fix bugs, and to make improvements – see Rossi and Bonaccorsi (2005). Without open source, they would have to do this independently, which would imply duplicated costs.

Empirical research in this area is at a nascent stage. Using the data at SourceForge, Fershtman and Gandal (2011) find empirical support for the existence of knowledge spillovers among open source projects. The paper shows that the structure of the project network is associated with project success and that there is a positive association between project closeness centrality and project success. This suggests the existence of both direct and indirect project knowledge spillovers among open source software projects.

7. Open Source More Broadly Defined – Digital Content[17]

Open source has spread well beyond the software development. Digital content is one area where open source has made major impacts. Creative Commons, which developed a way to help creators of content grant various degrees of copyright permissions to their work, is one of the most important outgrowths of open source. The Creative Commons licenses enable those who develop content to choose among a range of copyright protection, from "all rights reserved" (full protection) to “some

[17] This section draws heavily from comments and suggestions made by Greenstein.

rights reserved," to "no rights preserved." Several key institutions use Creative Commons licenses. Wikipedia,[18] the incredibly successful online encyclopedia, started with a variant of a GPL license for text, and then adopted 'Creative Commons' methodology. Some YouTube and Flickr users share their content using Creative Commons licenses. The success of Wikipedia[19] and other digital content providers using open source methodology shows that the open source model continues to evolve and will likely continue to be an important part of the digital economy.

References:

Athey, S., and G. Ellison, (2009) "Dynamics of Open Source Movements," mimeo.

Bessen, J., and E. Maskin, "Sequential Innovation, Patents, and Innovation," available http://econpapers.repec.org/paper/clanajeco/321307000000000021.htm

Boldrin, M. and D. K. Levine (2009), "Market Structure and Property Rights in Open Source," *Washington University Journal of Law & Policy*, PAGES

Bonaccorsi, A., and C. Rossi (2002), "Licensing Schemes in the Production and Distribution of Open Source Software: An Empirical Investigation" available at http://opensource.mit.edu/papers/bnaccorsirossilicense.pdf.

Bradner and Scott (1999), "The Internet Engineering Task Force," in DiBona, Chris, Sam Ockman, Mark Stone, 1999, Open Sources: Voices from the Revolution, O'Reilly Media Inc., Sebastopol, CA.

Chakravarty, S. Haruvy, E., Wu, F., 2007, "The Link between Incentives and Product Performance in Open Source," Development: An Empirical Investigation," *Global Business and Economics Review*, 9:151-169.

Casadesus-Masanell, R., and P. Ghemawat (2006): "Dynamic Mixed Duopoly: A Model Motivated by Linux vs. Windows." Management Science, 52 (7), 1072-1084.

Comino, S., Manenti, F., Parisi, M., (2007), "*On the Success of Open Source Projects*," Research Policy, 36: 1575-1586

DiBona, Chris, Sam Ockman, Mark Stone (1999), Open Sources: Voices from the Revolution, O'Reilly Media Inc., Sebastopol, CA.

[18] The 'wiki' concept was developed in 1995 by a software engineer named Ward Cunningham. Wikis were developed in order to fix bugs in software development, but now are applied to many other applications. See Greenstein (2011.)

[19] Wikipedia recently celebrated its tenth birthday. According to the Economist, Wikipedia now has over 17 million articles (3.5 million in English). The content is created and edited by users. It was ranked as the Internet's top research site in 2005, and consistently has been and continues to be one of the most popular websites. Currently it is used by a staggering 400 million users each month. See "Wiki birthday to you – a celebration of an astonishing achievement, and a few worries," the Economist, 13 January 2011.

DiBona, Chris, Danese Cooper, Mark Stone (2006), Open Sources 2.0: The Continuing Revolution, O'Reilly Media Inc., Sebastopol, CA.

Economides, N., and E. Katsamakas (2006): "Two-Sided Competitions of Proprietary vs. Open Source Technology Platforms and Implications for the Software Industry," *Management Science*, 52 (7), 1057-1071.

Fershtman, C., and N. Gandal, (2007), "Open Source Software: Motivation and Restrictive Licensing," *International Economics and Economic Policy*, 4: 209-225

Fershtman, C., and N. Gandal, (2011), "Direct and Indirect Knowledge Spillovers: The 'Social Network' of Open Source Projects," *RAND Journal of Economics*, 42: 70–91.

Ghosh, R., Glott, R., Krieger, B., and G. Robles (2002), "Free/Libre and Open Source Software: Survey and Study, Part IV: Survey of Developers," FLOSS Report, International Institute of Infonomics, University of Maastricht, The Netherlands, available at http://www.infonomics.nl/FLOSS/report/Final4.htm.

Greenstein, Shane, "Innovative Conduct in U.S. Commercial Computing and Internet Markets," Handbook on the Economics of Innovation, edited by Bronwyn Hall and Nathan Rosenberg, forthcoming, 2011.

Hann, I., Roberts, J., Slaughter, S., and R. Fielding, 2004, An Empirical Analysis of Economic Returns to Open Source Participation, available at http://wwwrcf.usc.edu/hann/publications_files/economic_returns_to_open_source_participation.pdf

Harhoff, D., J. Henkel, and E. von Hippel (2003), "Profiting from voluntary spillovers: How users benefit by freely revealing their innovations, *Research Policy* 32: 1753-1769.

Hars, A., and S. Ou (2001), "Working for free? - Motivations for participating in open source projects," *International Journal of Electronic Commerce*, 6: 25-39

Hertel, G., Niedner, S. and S. Herrmann (2003), "Motivation of software developers in open source projects: An internet-based survey of contributors to the Linux kernel," *Research Policy*, 32, 1159-1177.

Johnson, J. (2002), "Open Source Software: Private Provision of a Public Good," *Journal of Economics & Management Strategy*, 11: 637-662.

Johnson, J. (2006), "Collaboration, Peer Review and Open Source Software" *Information Economics and Policy*, 18: 477-97

Koski, H., "OSS Production and Licensing Strategies of Software Firms," *Review of Economic Research on Copyright Issues*, 2005.

Lakhani, K., and R. Wolf (2005), "Why Hackers Do What They Do: Understanding Motivation and Efforst in Free Open Source Projects, In: Feller/Open, J. Fitzgerland, S. Hissam, K. Lakhani (eds.), Perspectives on Free and Open Source Software, MIT Press, Cambridge.

Llanes, G., and R. De Elejalde, (2009) "Industry Equilibrium with Open Source and Proprietary firms," mimeo, available at http://www.hbs.edu/research/pdf/09-149.pdf.

Lerner, J., Parag, P., and J. Tirole (2006), "The Dynamics of Open-Source Contributors," *American Economic Review Papers and Proceedings*, 96:114-18

Lerner, J., and M. Schankerman, the Comingled Code: Open Source and Economic Development, MIT Press, Cambridge, MA, 2011.

Lerner, J., and J. Tirole (2002), "Some Simple Economics of Open Source" *Journal of Industrial Economics*, 52: 197-234

Lerner and Tirole, (2005a), "The Economics of Technology Sharing: Open Source and Beyond," *Journal of Economic Perspectives*, 19: 99-120.

Lerner, J., Tirole, J., (2005b)"The Scope of Open Source Licensing," *Journal of Law, Economics, and Organization*, 21: 20-56

Maurer, S. and S, Scotchmer, 2006, "Open Source Software: The New Intellectual Property Paradigm," *Economics and Information Systems*, pp. 285-322

Raymond, E. (2000), "The Cathedral and the Bazaar", available at http://www.catb.org /~esr/writings/cathedral-bazaar/cathedral-bazaar/

Rossi, C. and A. Bonaccorsi, 2005, Intrinsic vs. Extrinsic Incentives in Profit-Oriented Firms Supplying Open Source Products and Services, First Monday, Issue 10, No. 5, available at
http://pear.accc.uic.edu/htbin/cgiwrap/bin/ojs/index.php/fm/article/viewArticle/1242

Stallman, R., (1999), "The GNU Operating system and the Free Software Movement," in Dibona, C., Ockman, S., and M. Stone editors, *Open Sources: Voices from the Open Source Movement,* O'Reilly, Sepastopol, California.

von Krogh, G., and E. von Hippel, 2006, "The Promise of Research on Open Source Software," *Management Science*, 52:975-983.

West, J., and S. Gallagher, "Challenges of Open Innovation: The Paradox of Firm Investment in Open Source Software," *R&D Management*, 36: 315-328.

West, J., and S. O'Mahoney (2008), "The Role of Participation Architectures in Growing Sponsored Open Source Communities," *Industry and Innovation*, V 15 (2), pp. 145-168.