



The Comingled Code

**Open Source
and Economic
Development**

JOSH LERNER AND MARK SCHANKERMAN

The Comingled Code

Open Source and Economic Development

Josh Lerner and Mark Schankerman



The MIT Press
Cambridge, Massachusetts
London, England

© 2010 Massachusetts Institute of Technology

All rights reserved. No part of this book may be reproduced in any form by any electronic or mechanical means (including photocopying, recording, or information storage and retrieval) without permission in writing from the publisher.

For information about special quantity discounts, please email special_sales@mitpress.mit.edu

This book was set in Palatino on 3B2 by Asco Typesetters, Hong Kong.
Printed and bound in the United States of America.

Library of Congress Cataloging-in-Publication Data

Lerner, Joshua.

The comingled code : open source and economic development / Josh Lerner and Mark Schankerman.

p. cm.

Includes bibliographical references and index.

ISBN 978-0-262-01463-2 (hardcover : alk. paper) 1. Open source software. 2. Computer software—Development. I. Schankerman, Mark. II. Title.

QA76.76.S46L46 2010

005.3—dc22

2010001730

10 9 8 7 6 5 4 3 2 1

The Comingled Code

To Carol and Ralph

To my mother Etta and my father Paul (Z"L), with love

Preface and Acknowledgments

Many a book has a tangled story behind it, and this one has a particularly long and twisted one. As a result there are a lot of people to thank.

This project has sought to understand the role of open source in economic development using several approaches. Given the early stage of open source software's development and the inherent difficulty of measuring these activities, only through taking multiple approaches can we do justice to this complex phenomenon.

The first approach was a careful review of the economic principles to shed light on the open source phenomenon and its implications for economic development. We prepared an analytical framework that highlighted how economics could guide us in understanding these complex issues. To do so, we drew on a diverse array of bodies of literature, including work on growth, the nature of innovation, and the literature on incentives and innovation in open source. When economic principles point in different directions—as they sometime do—we highlighted the open questions. The analytic framework underlies the work, and broadly informs how we approach the conceptualization of the demand and the supply sides in open source, as well as how we address the key policy questions.

The second approach was the preparation of a series of case studies of the open source phenomenon. We sought to understand the complexity of the role of open source in half a dozen nations in various stages of development. Rather than focus on the entire canvas of activities in each nation, we looked at different issues raised by specific situations in a number of emerging economies. Each study was based on interviews with Harvard Business School's case writers and secondary sources. The cases we studied were:¹

1. A variety of supplemental material not included in the book is posted on-line at <http://mitpress.mit.edu/comingledcode>.

- In Brazil, a project called HackerTeen seeks to combine educating young programmers with the development of new software.
- In China, the CEO of a mobile telephone software company must decide whether to use Linux or commercial alternatives as the basis for his system.
- In France, a chief information officer must choose whether to recommend the Ministry of Finance will run on open source or proprietary software.
- In Singapore, the government considers whether to support a research initiative to promote research in open source.
- In South Africa, a vendor of software and services must decide how much effort to devote to developing expertise in open source.
- In Thailand, the government considers whether to promote a “People’s PC,” which would combine a low-cost hardware platform with open source software.

By focusing on a wide variety of actors in a diverse array of situations, we gained a rich picture of open source activity and its implications for development.

The third and final aspect of the project involved a large-scale survey of software users and developers in fifteen countries. (The developer survey alone had nearly 2,000 respondents.) The nations included many of the same ones on which we wrote cases, as well as other developing and industrialized countries, including Chile, Greece, India, Israel, Kenya, Mexico, Poland, Russia, and Turkey. We highlighted in the questionnaire not just questions about the utilization of open source software, but also about the respondents’ attitudes and perspectives on the costs and benefits of this software. We further looked separately at developers and users of different types (for users, different sizes of companies, government agencies, and ownership; for developers, different sizes, ownership, software activities, etc.).

After we had conceptualized the project, we realized that the expense was so great that it was unlikely that we could fund it using our own resources: both the surveys and case studies proved to be extremely costly to implement. We discovered that Microsoft was interested in funding academic work into open source, with the goal of promoting a less ideological discussion of the pros and cons of software choices and public policy toward this sector, as well as providing

the empirical evidence that could contribute to a more balanced and evidence-based formulation of public policy. We accepted their funding under stringent terms that ensured that the effort was characterized by intellectual independence and analytical rigor.

This work, and the conversations it engendered, led to the idea of developing this book. As always, the process of converting research into (at least what we hope is) readable prose proved to be a far more daunting task than we had initially envisioned.

First of all, we are very grateful for the contributions of Jacques Crémer of the University of Toulouse. In addition to being the co-author of chapter 6, Jacques was crucially involved in many discussions that led to the development and elaboration of the main themes and ideas of the book. His was an important contribution to the volume.

Next, our thanks must go to our long-suffering MIT Press acquiring editor, Jane Macdonald, for her patience with us. We also would like to thank David Evans, Anne Layne-Farrar, and Daniel Schwartz of LECG for their initial encouragement and help with the project, especially in the development of the survey questionnaires. Jacques Lawarree was the project's champion at Microsoft. We also benefited from considerable financial support from Harvard Business School's Division of Research. Mark Schankerman also thanks the British Academy and the Muzzy Chair in Entrepreneurship at the University of Arizona for their financial support of this research.

Sam Kortum of the University of Chicago was a member of the original project team and contributed a number of ideas. Brian DeLacey, Kerry Herman, and David Kiron played key roles in developing the case studies. Ivan Maryanchuk and Liat Oren provided excellent research assistance on the empirical analysis found in chapters 4, 5, and 6, and Kathy Han and Gabriel Fotsing provided invaluable research support on other aspects of the book. Maurie SuDock helped with the manuscript preparation and cleanup. While we are grateful for all the support, it is important to note that the ideas and recommendations represent our own opinions only.

As we discuss at more length in the introduction, there remains a considerable divide between the economics and open source communities. It is our hope that this book helps bridge this gap, by discussing the economic issues around open source in an accessible but tangible way based on systematic empirical evidence. We particularly hope

that this approach leads to a less shrill (or less acrimonious) interaction between these communities, and a more informed, evidence-based public policy toward this sector.

Josh Lerner and Mark Schankerman
Boston and London
October 2009

Contents

	Preface and Acknowledgments	ix
1	Introduction	1
2	Software and Growth	15
3	The History of Open Source	35
4	The Supply Side: Comingling Open Source and Proprietary Software	61
5	The Demand Side: Assessing Trade-offs and Making Choices	103
6	Assessing Government Policies toward Software (with Jacques Crémer)	157
7	The Takeaways	207
	Glossary	215
	References	225
	Index	231

1 Introduction

Open source software involves developers at many different locations and organizations sharing code to develop and refine computer programs that are then distributed at no or low direct cost. Over the past fifteen years open source software has experienced explosive growth around the world. The importance of open source software today can be illustrated by considering a few examples:

- The market for server software, which is used by the computers that make Web pages available to users through the Internet, has been dominated by the open source Apache project since the inception of systematic tracking by Netcraft in 1995. As of July 2009 more than half of servers employed Apache or other open source products, rather than commercial alternatives from Microsoft, Oracle, or other firms.¹
- While definitive numbers are very hard to come by, a variety of survey evidence suggests that the open source operating system called Linux has rapidly outstripped Microsoft's Windows program as the operating system most frequently embedded into products ranging from mobile phones to video recording devices.²
- Open source software is dominant in a number of other areas as well; for example, PERL and PHP are the dominant scripting languages (programming languages designed specifically for instructing one computer how to communicate with others).
- Even corporations that have traditionally resisted open source, and the stringent General Public License (GPL) in particular, have seemingly changed their approach. Most visibly Microsoft, whose executives

1. http://news.netcraft.com/archives/web_server_survey.html (accessed August 23, 2009).

2. <http://www.linuxfordevices.com/c/a/Linux-For-Devices-Articles/Snapshot-of-the-embedded-Linux-market-April-2007/> (accessed August 23, 2009).

once branded the GPL as “fundamentally undermin[ing] the independent commercial software sector,”³ released two substantial blocks of code under this license in July 2009.⁴

- Open source software is not a phenomenon that is confined to rich countries. For better or for worse, the Brazilian and many other developing nation governments are promoting the use of open source software as an alternative to proprietary products. Significant numbers of contributors to open source software, in proportion to the population, can be found in countries with per capita income as low as \$10,000.

Open source software may be poised for rapid growth in the future. The number of projects has exploded: the website SourceForge.net, which provides free services to open source software developers, has grown from a handful of projects in 2000 to well over two hundred thousand open source projects today.⁵ Many of the projects seem to have room to expand: for instance, the operating system Linux has opportunities in the market for desktop operating systems; in 2009, only one percent of the Web queries tracked by Net Applications came from machines running Linux, although that share was gradually rising.⁶ More generally, the economic downturn appears to have accelerated corporate interest in and adoption of open source solutions: for instance, IDC recently revised its projected growth in revenue from open source products through 2013 upward, to an annual rate of 22.4 percent.⁷

The growth of open source software is attracting considerable attention from the public sector as well. Government commissions and agencies have proposed—and in some cases implemented—a variety of measures to encourage open source developers, including R&D support, encouragement for open source adoption, explicit preferences in government procurement, and even mandates regarding software

3. <http://www.microsoft.com/presspass/exec/craig/05-03sharedSource.msp> (accessed August 23, 2009).

4. http://www.microsoft.com/presspass/features/2009/Jul09/07-20LinuxQA.msp?rss_fdn=Toppercent20Stories (accessed August 23, 2009). It should be noted that Microsoft’s motivations for this step were hotly questioned and debated (e.g., <http://blog.seattlepi.com/microsoft/archives/174828.asp>; accessed August 23, 2009).

5. <http://sourceforge.net/apps/trac/sourceforge/wiki/Whatpercent20ispercent20SourceForge.net?> (accessed August 25, 2009).

6. <http://marketshare.hitslink.com/report.aspx?qprid=8> (accessed August 23, 2009).

7. <http://finance.yahoo.com/news/Open-Source-Software-Market-bw-400190557.html?x=0&.v=1> (accessed August 23, 2009).

choices. In 2008 the Center for Strategic and International Studies identified 275 open source public policy initiatives, 182 of which have been favorably enacted.⁸ For instance, since 2003, Singapore has offered tax breaks to companies using GNU/Linux operating systems rather than proprietary ones in order to encourage the development of the local software sector. Many European governments have enacted policies to encourage the use and purchase of open source software for government use. Governments have even mandated the development of localized open source projects, as has occurred in China. But while the efforts are concentrated in Europe and Asia, the interest in open source is truly global.

Brazil, Mexico, and South Africa are just a handful of developing nations that have launched significant open source initiatives. In part the appeal of these programs is that they are typically available for free, or at a much lower direct cost than comparable proprietary products. But these nations also argue that a wide variety of economic development benefits can follow from the development of a vibrant open source community, including the development of local industries, an improved foreign trade balance, and a reduction in intellectual property piracy.

Policy discussions around open source, though, have frequently been characterized by more heat than light. While the question at the heart of the debate—what is the impact of open source on consumers, firms, and economic growth more generally—is an economic one, the discussion is frequently framed in nearly theological tones. Advocates of open source passionately assert its benefits, while critics denigrate its role. Missing from the debate has been rigorous economic analysis and systematic microeconomic evidence, which might help sort out these competing claims.

This paucity of rigorous analysis in part reflects the strong emotions that the subject engenders. But it also reflects the difficulty of conclusively answering these questions empirically. Open source software usage is difficult to track, and the economic impacts of software utilization hard to trace definitively. Moreover there is always the chicken-and-egg problem: Did the use of a certain type of software promote the growth of a software industry in a given country, or did the rapid growth of a nation lead to it turning to a given set of software?

8. http://csis.org/files/media/csis/pubs/0807218_government_opensource_policies.pdf (accessed August 23, 2009). All citations to these programs are provided in chapter 6.

About This Book

This book seeks to address these challenging issues. Building on a series of analyses, we hope to improve our understanding of how open source and proprietary software interact and the policy issues that this raises. In particular, we try to provide an economic perspective on a debate that has been largely conducted on other terms, and a new large-scale database that can support more informed discussion.

The prior paragraph, of course, suggests a question: Why should economists have anything to say about software at all? It is true that we have spent many a happy hour programming regressions in Stata and trying to make tables come out just right in LaTeX. But we certainly do not pretend to be expert programmers, nor have we ever organized an open source project or corporate software initiative.

Rather, the answer to this question lies elsewhere. The question might well be reasonable if the book's focus was on the nature of open source and/or proprietary programs. But our focus in this book is different. In particular, we will be concentrating on understanding the broader impact of open source software. And once we move beyond the programs themselves, and to how they interact with the rest of the world, an economic perspective can be indispensable. Among the questions we will examine are these four:

- How does software differ from other technologies when it comes to promoting economic development?
- What are the motivations that drive individuals and firms to contribute to open source projects?
- How do firms using and developing software view the trade-offs between proprietary and open source projects?
- What policies can governments adapt to ensure that open source effectively competes with proprietary software and contributes to economic growth, and which steps run the danger of backfiring?

It is our belief that in each of these arenas, an economic perspective adds essential value in shedding light on these questions.

In particular, we highlight several crucial insights that the application of an economic framework to the world of open source suggests:

- Traditional economic frameworks that prescribe that the market will solve the optimal allocation of activity do not apply very well to the software industry.

- Open source and proprietary software share many common elements, as well as differences, that economic analyses can illustrate.
- Based on studies completed to date, it is hard to draw unambiguous conclusions as to the superiority of open source versus proprietary software.
- There is not a strong foundation for the claim that the government should favor open source or proprietary software when purchasing for its own purposes.
- While governments in developing countries may have strong rationales for encouraging the development of a software industry in general, the desirability of encouraging open source is more circumscribed, and economic analysis help identify the conditions under which such support may be justified.

Ultimately, it is our hope that this book stimulates more conversations between the open source and economics communities. Each side, it seems, has much that can be learned from the other. Despite (perhaps because of) the extensive research that we have undertaken in this area, we are keenly aware of the limitations that economists have faced when studying the open source realm and the way in which a keener understanding of the technical and social aspects of open source communities could be valuable. At the same time we strongly believe, and hope to have shown in this book, that many of the conceptual and methodological tools in the economics “tool kit” can be valuable additions to many in the programming community interested in more rigorous, systematic studies of this important sector of the economy.

A Road Map to the Book

To give a sense of where we are going with the book, we will provide a brief overview of chapters 2 through 7.

An Initial Look at Software and Growth

The second chapter has an ambitious mandate. To help set the stage for an understanding of open source software, we begin by explaining why software is important from an economic perspective.

Traditionally economists studying economic growth focused on how physical capital accumulates over time. As the economy grows and people take some of their earnings and invest it in extra machines

and facilities, there is greater output and hence economic growth. Later theories applied a similar principle to the accumulation of “human capital,” where savings are used to accumulate more knowledge: some of labor time is used not directly for production but for schooling and training.

In recent years, however, there has been a dramatic change in the economists’ perspective, growing out of the realization that we had “Hamlet without the prince.” The “new growth theory” puts technological innovation at the center of the growth process: the ways that inputs (e.g., people and machines) are translated into outputs, both products and services. The special thing about an innovation is that it can be shared: everyone can use it at the same time (whereas a worker or a machine cannot be everywhere at once). Unlike a physical piece of equipment, the use of a recipe for a better mosquito repellent in one country or by one individual does not hinder its use in other countries or by other individuals. Economists call this the nonrivalry property of information. But one must also ensure that individuals and firms have incentives to invent better recipes. They might not have incentives to do so if the new knowledge they generate is immediately and freely made available to everyone who wants to use it.

Where does computer software, and open source software in particular, fit within this framework of growth theory? In many respects it is like other technologies: it allows people to produce more with the same amount of materials. Indeed software’s reach extends far beyond the software industry. Much of the innovation in software sprang from firms in other industries that have embodied software into products and processes.

The same conundrum described in the new growth theory is present in traditional proprietary software. On the one hand, since a software program, once developed, can be distributed very cheaply via the Internet, it is socially wasteful for some firms and countries to use inferior software. And indeed software use (at least as evidenced by available evidence on diffusion of Internet and computers) and software production (measured using software patenting in the United States and by open source contributions) both vary with the level of development: many national information technology industries appear stuck at a lower level of development (these disparities appear to be much greater on the production side). On the other hand, as before, we need to worry about the incentives to develop better software.

Where does open source software fit within this framework? If the open source model of software development could deliver state-of-

the-art and easy-to-use products for all applications, it would solve the conundrum of the new growth theory as it applies to software. On the one hand, it would make the best recipes available everywhere at essentially zero cost, hence taking full advantage of the nonrivalry property. On the other hand, the incentive problem would be solved by the fact that programmers (either as individuals or as firms) contribute voluntarily to the software development. This analysis suggests why open source is potentially so revolutionary, and the critical importance in understanding its economic development impact.

Lessons from History

To understand open source software, it is useful to understand where it has come from. In chapter 3 we explore the origins and evolution of this sector. We highlight how there have been three distinct eras of open source development:

- During the 1960s and 1970s many of the key features of computer operating systems and the Internet were developed in academic settings such as Berkeley and MIT, as well as in central corporate research facilities where researchers had a great deal of autonomy, such as Bell Labs and Xerox's Palo Alto Research Center. The sharing by programmers in different organizations of the source code for computer operating systems and for widely used transmission protocols was commonplace. These cooperative software development projects were undertaken on a highly informal basis.
- In response to the threats of litigation engendered by this lack of clear rules, efforts to formalize the ground rules behind the cooperative software development process emerged. These efforts ushered in the second era. The critical institution during this period was the Free Software Foundation, begun by Richard Stallman of the MIT Artificial Intelligence Laboratory in 1983. The foundation sought to develop and disseminate a wide variety of software at no cost. The Free Software Foundation introduced a formal licensing procedure, called a General Public License, for a computer operating system called GNU. This procedure ensured that all derivatives of the program would also be disseminated at low or no cost.
- The widespread diffusion of Internet access in the early 1990s led to the third era that saw a dramatic acceleration of open source activity. Because it became easier for developers at very distant physical locations to collaborate on the development of projects at low cost, the volume of contributions and diversity of contributors expanded sharply,