

NBER WORKING PAPER SERIES

PATENT COMMONS, THICKETS, AND OPEN SOURCE SOFTWARE ENTRY
BY START-UP FIRMS

Wen Wen
Marco Ceccagnoli
Chris Forman

Working Paper 19394
<http://www.nber.org/papers/w19394>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 2013

We thank Ajay Agrawal, Ashish Arora, Tim Bresnahan, Alberto Galasso, Stuart Graham, Patrick Legros, Josh Lerner, Megan MacGarvie, Marc Rysman, Timothy Simcoe, Salvatore Torrisi, Nicolas Van Zeebrock, Feng Zhu, and participants at the NBER Patents, Standards, and Innovation Conference, the 2011 "Workshop on Information Systems and Economics (WISE), the 2011 Conference on Information Systems and Technology (CIST), the 2011 REER conference at the Georgia Institute of Technology, the 2012 "European Policy for Intellectual Property (EPIP) conference, the Université Libre de Bruxelles (ECARES), the 2012 Academy of Management Conference Symposium, the 2013 Platform Strategy Research Symposium at the Boston University. We gratefully acknowledge funding from a Kauffman Foundation/Georgia Research Alliance grant for the study of entrepreneurship and productivity. We thank Nirmalya Choudhury, Matthew Espy, Emily Getreu, Bridget M. Gorta, Sujay Mehta, Daniel Mitchell, and Jian Zhao for outstanding research assistance. Wen Wen thanks the Kauffman Foundation for providing funding for this research through a Kauffman Dissertation Fellowship. Chris Forman acknowledges funding from the Alfred P. Sloan Foundation through an Industry Studies Fellowship. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2013 by Wen Wen, Marco Ceccagnoli, and Chris Forman. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Patent Commons, Thickets, and Open Source Software Entry by Start-Up Firms
Wen Wen, Marco Ceccagnoli, and Chris Forman
NBER Working Paper No. 19394
August 2013
JEL No. L86,O34

ABSTRACT

We examine whether the introduction of a patent commons, a special type of royalty free patent pool available to the open source software (OSS) community influences new OSS product entry by start-up software firms. In particular, we analyze the impact of The Commons—established by the Open Source Development Labs and IBM in 2005. We find that increases in the size of The Commons related to a software market increase the rate of entry in the market by start-ups using a new product based on an OSS license. The marginal impact of The Commons on OSS entry is increasing in the cumulativeness of innovation in the market and the extent to which patent ownership in the market is concentrated.

Wen Wen
The University of Texas at Austin
McCombs School of Business
2110 Speedway Stop B6500
Austin, TX 78712
Wen.Wen@mcombs.utexas.edu

Chris Forman
Georgia Institute of Technology
Scheller College of Business
800 West Peachtree Street, NW
Atlanta, GA 30308
chris.forman@scheller.gatech.edu

Marco Ceccagnoli
Scheller College of Business
Georgia Institute of Technology
800 West Peachtree Street, NW
Atlanta, GA 30308
marco.ceccagnoli@scheller.gatech.edu

1. Introduction

While patents play an important “property rights” role in facilitating transactions in markets for technology (e.g., Arora et al. 2001, Gans et al. 2002), the strategic use of patents and the appearance of dense, overlapping webs of property rights known as patent thickets may also work to stifle innovation (e.g., Bessen and Meurer 2008, Jaffe and Lerner 2004, Shapiro 2001). Empirical evidence on this issue is still mixed, with implications varying across industries and firms.¹ The presence of patent thickets may be a particular issue for developers of information technology (IT) products like computing hardware and software, where innovation is often highly cumulative and products rely on standards of heterogeneous inventions for whose patent rights are often owned by many different firms.² Small firms often find navigating such patent thickets difficult, as they will typically not have a large intellectual property rights (IPR) portfolio that they can use to cross-license with existing IPR holders.

A range of institutional mechanisms have appeared to help companies navigate patent thickets, including standard setting organizations (SSOs) and patent pools. Some of these mechanisms will lower the transaction costs of identifying and negotiating licensing agreements for related technologies, but may also increase the incentives for some patentees to litigate (e.g., Lampe and Moser 2013, Simcoe et al. 2009). As a result, they have had mixed effects on inventive activity that builds upon earlier patented technologies.³

An alternative approach is to offer IPR through a patent commons, a special type of patent pool that offers royalty-free usage of patents to any firm that promises not to assert their IPR against the commons’ beneficiaries (Hall and Helmers 2013, L  v  que and M  n  re 2007, Serafino 2007). While the number and size of such donations have been growing rapidly, thus far we have little evidence on whether they in fact contribute to increased innovative activity.⁴ This is an important gap in understanding. The literature suggests that inventors may have insufficient incentives to provide intellectual property as a public good to the community (Gambardella and Hall 2006), so patent commons may contain few or low-

¹ For example, Hall and Ziedonis (2001) study on the semiconductor industry finds that the pro-patent shift in U.S. policy in the 1980s spawned an increase in strategic patenting among capital-intensive firms, but also facilitated entry by specialized design firms. Cockburn and MacGarvie (2011) find that, following expansions in the patentability of software in the mid-1990s, average entry rates declined but the likelihood of entry by firms holding patents increased.

² For example, Biddle et al. (2012) identify 251 technical interoperability standards in a modern laptop. Of these 251 standards, 112 were developed by consortia, 90 by formal standards development organizations, and 49 were developed by individual companies.

³ For example, Rysman and Simcoe (2008) show that citations to patents increase significantly after disclosure that they are part of a standard, while Lampe and Moser (2010) study the sewing machine patent pool and find that the pool decreased patenting and innovation, particularly among members of the pool.

⁴ One recent exception is Hall and Helmers (2013), whose recent study of the “Eco-Patent Commons” showed that “making patents accessible royalty-free did not result in any significant increase in diffusion as measured by citing patents” (Hall and Helmers 2013). However, their results must be interpreted with caution, given the short period of time following the establishment of this particular patent commons in 2008.

quality patents that have little impact on innovative activity.⁵ In addition to the potential low quality issue associated with poor incentives, the patents pledged in the commons may not contain all of the technologies necessary to produce and commercialize a new invention. If other necessary technologies are not held within the commons, barriers to commercialize the invention may still be high. In short, our research speaks to the efficacy of a growing mechanism to help firms navigate patent thickets.

Motivated by these observations, we take a first step to evaluate how the creation of a patent commons affects innovation in the open source software (OSS) setting. In particular, we examine implications of the introduction of The Patent Commons (commonly referred to as “The Commons”), initiated by the Open Source Development Labs and IBM in 2005. While innovation can be examined from many perspectives, in this study we focus on start-up entry into a market through new product introductions. We show that introduction of The Commons influences new product entry by start-ups using an OSS license (which we refer to as “OSS entry”). Our key argument is that because The Commons could potentially mitigate licensing cost and litigation threats caused by patent thickets, start-ups that have access to it would expect lower sunk costs of entry than otherwise and thus be more likely to enter into downstream markets. Using a similar logic, we argue that such institutional mechanisms will be particularly valuable in environments with high licensing and litigation costs. We highlight the following two settings where such costs would be the highest and thus The Commons will have the greatest impact on OSS entry—when innovations are highly cumulative and when IPR ownership (particularly patent ownership) is very concentrated.

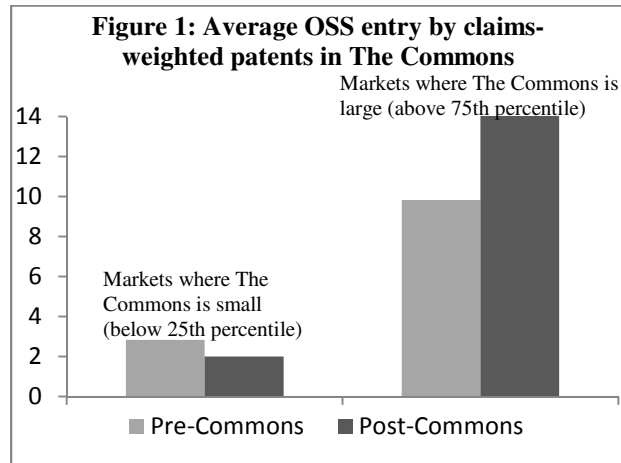
To test these hypotheses, we assemble data on OSS entry using longitudinal press releases from 2,054 U.S. start-up software firms contained in the Gale database “PROMT”. Using count data conditional fixed effects models, our empirical strategy examines whether time series variation in the number of patents contributed to The Commons related to a narrowly defined software market is associated with changes in the amount of OSS entry into that market. A key and challenging task in assembling our dataset was to identify patents relevant to entry in product markets. Following prior work that has examined the extent to which patents deter entry into the software industry (Cockburn and MacGarvie 2009, 2011), we allocate patents to software product markets based upon the technological classes of patents and their key words.

Our results show that a 10% increase in The Commons’ patent claims in a software market is associated with a 1.3%-2.9% increase in the rate of OSS entry by start-ups into that market.⁶ However,

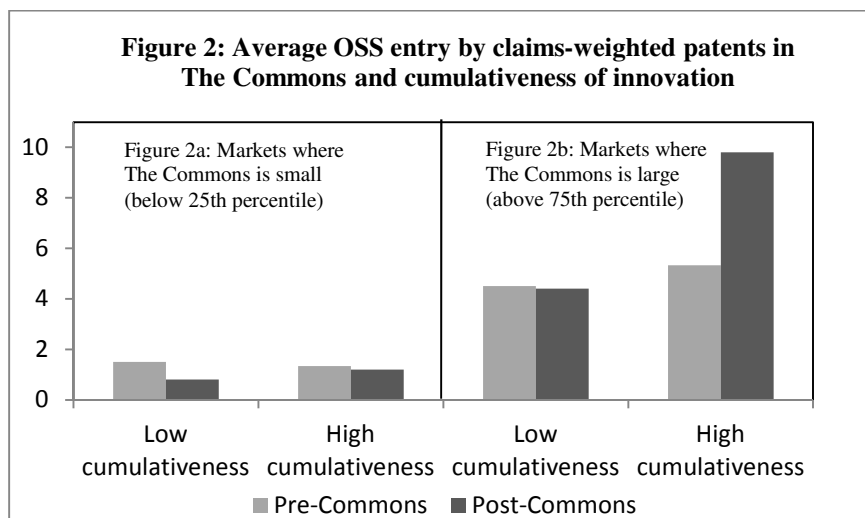
⁵ While we do not explicitly focus our research on the question of why firms may contribute some of their patents to the commons, we do discuss IBM’s incentives for establishing a patent commons in 2005 to motivate our analysis as well as to probe our identification strategy (see sections 3.2 and 6.3).

⁶ All marginal effects are evaluated at mean values of covariates.

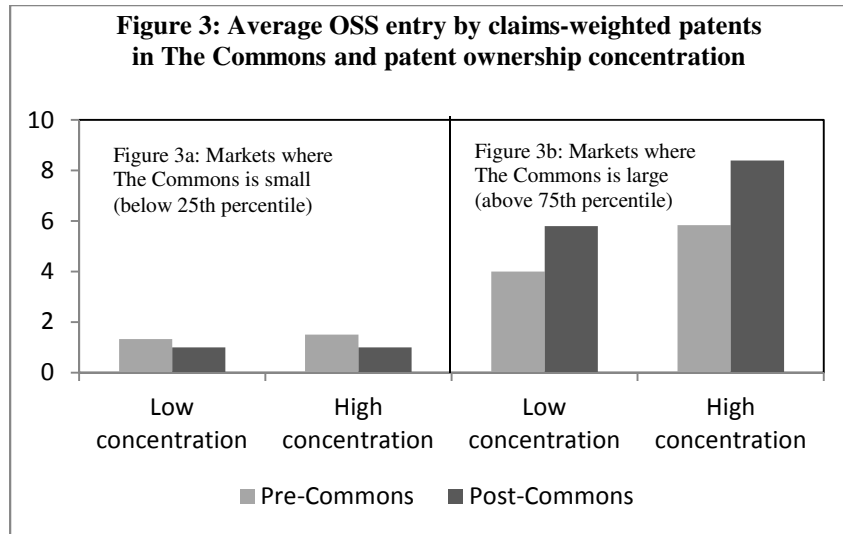
introduction of The Commons influences entry only in those markets where innovations are highly cumulative and when patent ownership is concentrated. Figures 1, 2 and 3 foreshadow our main results. Figure 1 shows that there is significant OSS entry after introduction of The Commons in markets where the number of claims-weighted patents in The Commons is large, while there is no change in OSS entry in other markets. Figures 2 and 3 motivate our results demonstrating differential benefits for The Commons in markets with high cumulateness and concentration. Figures 2a and 3a show that there is little growth in OSS entry after the introduction of The Commons in all markets where The Commons has few patents. Figures 2b and 3b examine growth in entry in markets where The Commons is well represented; Figure 2b shows that the introduction of The Commons is associated with greater OSS entry but only in markets where innovation is cumulative, while Figure 3b shows that growth in entry post-Commons is higher in markets with high IPR concentration.



Notes: For markets with large size of The Commons, the t-test statistic for the mean difference between pre-Commons and post-Commons OSS entry is 1.313 with a p-value 0.1.



Notes for Figure 2b: In markets with high cumulateness, the t-test statistic for the mean difference between pre-Commons and post-Commons OSS entry is 1.970 with a p-value 0.04.



Notes for Figure 3b: In markets with high concentration, the t-test statistic for the mean difference between pre-Commons and post-Commons OSS entry is 1.311 with a p-value 0.1.

A particular concern is that donations to The Commons could be correlated with unobserved market characteristics that vary over time. We examine the robustness of our results to a variety of changes to our estimation approach. We first show that our results are robust to controlling for market-specific time trends. We next implement a series of falsification exercises. We first examine the timing of the effects of The Commons. We find that growth in The Commons is not associated with increased entry from 1999 to 2003, but do find that it is associated with entry in 2004, one year prior to its introduction in 2005. We argue that that this finding reflects the implications of other activities that IBM was engaged in to promote open source over the same period, including public announcements of IBM's support of OSS. These activities were in turn prompted by the SCO's lawsuit against IBM in 2003, which increased the perceived legal risks to developers of OSS.⁷ As we discuss in further detail below, one of the benefits of The Commons to developers of OSS is as a signal of the commitment of IBM to provide support to the OSS community in the face of these legal uncertainties. We do not attempt to separately identify these benefits from others provided by The Commons, but instead acknowledge that The Commons will lower the expected legal costs of developing OSS through multiple channels. We further show that growth in The Commons is not associated with increased entry from products that should not see entry costs fall with its introduction, namely products offered under a proprietary license.

Last, we examine the robustness of our results to the use of instrumental variables. Our identification strategy uses variation across markets and time in the extent to which IBM holds valuable blocking patents, as proxied by IBM patents opposed at the European Patent Office (EPO). We also

⁷ On March 7, 2003 IBM was sued by the SCO Group, which asserted that IBM's contributions to Linux infringed on SCO's UNIX System V source code.

exploit an additional source of time variation associated with changes in the perceived risks to OSS that occurred during our sample period. Our results using these models are consistent with the baseline described above.

2. Theoretical Framework

2.1 How patent commons facilitates OSS entry

Patent thickets can increase at least three broad types of sunk costs of entry for start-ups: (1) the costs of inventing around existing patents; (2) the costs of infringement, which may include the costs of licensing the infringed technology and the costs of litigation such as acquiring a defensive patent portfolio as well as injunction and damages; and (3) the transaction costs of acquiring patents owned by others.

While firms that have a defensive patent portfolio may be able to navigate the patent thicket through cross-licensing their own patents with those of other IPR holders, this strategy will be harder to implement for small firms appropriating value from OSS. Firms that rely on the OSS community⁸ for innovation inputs need to conform to the norms of the OSS developers, who are often philosophically opposed to software patents, considering them antithetical to the spirit of freedom that imbues OSS development (Marson 2004, Schultz and Urban 2012, Stallman 2011).⁹ Further, the costs of writing and administering patents may be too high for small firms relative to their benefits. For example, they may have insufficient resources to hire legal staff.

As has been noted elsewhere, incumbent firms with significant patent portfolios and who use traditional appropriability mechanisms to commercialize new technologies in other settings will sometimes also contribute to open source communities under a so-called “private-collective model of innovation” (von Hippel and von Krogh 2003). Under this model, incumbent firms may contribute to public goods such as OSS. While they may be unable to appropriate value directly from the public good, they may be able to create and appropriate value from complementary products and services such as downstream application software that interfaces with open source or support for OSS.¹⁰ While such firms may be able to navigate patent thickets by using their own patent holdings in cross-licensing agreements, the inability of other small firms to similarly navigate the thicket may reduce the extent of complementary innovation and in so doing reduce the value of the public good.

⁸ We commonly refer to this type of firms as an OSS firm. More specifically, an OSS firm is defined as one that uses and commercializes source code that meets the Open Source Initiative definition of open source software.

⁹ OSS firms may even fear the existence of patents that are even pledged to the community for purely defensive purposes; while the initial motivates may be altruistic, it remains possible that the original patent holder may have a change in strategy, in particular a change in ownership or management may initiate a change in how patents are used within the organization (e.g., Schultz and Urban 2012).

¹⁰ For example, Fosfuri et al. (2008) find that firms will be more likely to produce OSS products when they have large stocks of complementary patents or downstream complementary capabilities.

One institutional mechanism that these incumbents have used to stimulate complementary innovation is to donate patents to a commons for use only within OSS community. Such donations go beyond the requirement that contributors to OSS products automatically grant a license to use, modify, and redistribute contributed code to all other legitimate users of the code but also extends to complementary technologies that may be used in conjunction with existing OSS products. Contributors not only incur the opportunity cost of foregone profits from not enforcing the IPR, but also bear an explicit cost of maintaining and renewing the patent. While the motivations to contribute to patent commons are in themselves interesting, it is not our goal to explore them, as they have been investigated elsewhere (Alexy and Reitzig 2013).

Contributions to patent commons can have direct and indirect benefits to the OSS community. For the start-up firms that we study (i.e. those that intend to enter into a market with OSS products), one direct benefit will be the reduction in expected invention and licensing costs related to the donated technology. Instead of inventing around, start-ups can directly use the technology described by the commons' patents. Moreover, start-ups can use the pledged patent to substitute for protected technologies that may block entry but are not in the commons, avoiding the costs of transacting and negotiating with patentees. These benefits will be increasing in the number of patents that are made available to the start-ups. Second, as incumbents increase their contributions it will signal their commitment to generate profits as a private-collective (p-c) innovator through the production of complementary goods and services, rather than through direct enforcement of IPRs. While the contributor retains the right to enforce other patents not included in the donation, doing so would harm its reputation to the OSS community, decreasing its ability to create and appropriate value as a p-c innovator. As a result, the perceived threat of litigation will decline, or, to use a widespread terminology among software industry practitioners, there will be less "*fear, uncertainty, and doubt*" (FUD, cf. Auza 2011).

There may be other indirect benefits. As Alexy and Reitzig (2013) describe, the donation of patents may encourage reciprocal behavior from other industry participants that may benefit p-c innovation. For example, other firms involved in p-c innovation may also choose to pledge patents for use by the community, and may acquire additional patent rights to prevent proprietary innovators from acquiring the same and using them against the community. As more patents in a technology area are pledged in the commons, it is more likely to create norms of non-enforcement and encourage other p-c innovators' contribution. One way to view this result is through the lens of public goods. The contribution of patents represents a commitment not to assert IPR with the goal of fostering the development of new software by entities who are unable to protect themselves against IPR enforcement by traditional means. This goal is more likely to be accomplished when contributions are made by multiple firms. However, in an asymmetric information environment, firms have an incentive to under-contribute (Coase 1960). To

provide the public good, actors with high valuations must contribute more than those with lower valuations (Mailath and Postlewaite 1990). In our setting, this corresponds to the initiator of the commons providing more than other firms, and so the likelihood of forming successful patent commons will be increasing in the initial contribution.

In sum, we expect that patent commons may effectively reduce licensing costs and litigation threats for start-ups that intend to enter into a market with new OSS products (i.e. OSS entry), and such benefits will be increasing in the size of the original pledge. It is also worth noting that these positive effects on OSS entry will be reinforced as firms enter and produce complementary OSS products, as these complementary products would create a platform of interlocking components, which increases the value of OSS products and services to potential buyers and thus further facilitate entry.

2.2 How the value of patent commons varies with market characteristics

In the previous section we showed how a patent commons can promote OSS entry by reducing expected invention costs or by alleviating infringement costs and transaction costs of negotiating with patentees. In this section we explore market circumstances where the effects of such a patent commons on reducing entry costs, and so encouraging entry, will be greatest.

A body of literature has attempted to quantitatively characterize patent thickets in specific technology areas and examined their implications for innovation and competition (Hall et al. 2012, von Graevenitz et al. 2011, Ziedonis 2004). We focus on two observable dimensions of patent thickets that influence the ex ante costs of entry. The first is the cumulateness of innovation, which has been defined as the extent to which an innovator builds on prior developments and discoveries (e.g., Green and Scotchmer 1995, Scotchmer 2004: 127). In an environment with high cumulateness, the boundaries of potential blocking patents are usually blurred. This makes it difficult to build upon existing patents and leads to high costs of inventing around. Further, the presence of many cumulative innovations also suggests high infringement costs and transaction costs, as start-ups could easily infringe and may need to obtain licenses for a large set of related patents to enter into a technology space.¹¹ Therefore, because patent commons could reduce these costs, its effects on OSS entry will be greatest in environments where these ex ante costs are highest, as is the case for environments with cumulative innovations.

The second market characteristic that we study is the concentration of patent ownership within a software market, defined as the extent to which patents are distributed across different holders. Two

¹¹ It is worth noting that the “cumulateness of innovation” in our context is different from the “royalty stacking” discussed by Lemley and Shapiro (2007). Royalty stacking refers to situations in which a single product potentially infringes on many patents, and thus the royalty rate to one patent holder is affected by the rates to the holders of other patents reading on that product. While high cumulateness also suggests an innovation is built upon many others, it does not focus on the interaction among the potential licensee and many patent holders, and the resulting royalty rate. We will focus on this in the next paragraph—how the ownership of existing patents faced by a start-up affects the costs of licensing negotiations.

views have recently been set forth about how patent ownership influences the costs of licensing negotiations. One view holds that increases in fragmentation of patent ownership (i.e., decreases in concentration of patent ownership) will increase the transaction costs of licensing patents, creating an “anti-commons” effect (Heller and Eisenberg 1998). Under this view, when there are many small exclusionary patents held by many firms, the transaction costs of coming to terms with many patent holders will influence a firm’s strategic response to potential expropriation risks. For example, Ziedonis (2004) shows that firms will patent more aggressively when patent ownership is fragmented. In sum, under this view, the transaction costs of acquiring patents owned by others are particularly high when patent ownership is fragmented.

Recent work has challenged the anti-commons view. Under this view, concentration of patent ownership increases the value of the negotiation for the patent holder, resulting in an increase in incentives to litigate (Galasso and Schankerman 2010, Lichtman 2006). For example, consider the case where patents are alternatively held by 1000 firms or 1 firm. While transaction costs are potentially higher in the former setting, the potential licensee may not need to obtain licenses from each of the 1000 firms as some patentees may not litigate the patent if there are non-zero costs to litigation and some uncertainty of whether the focal patent will be upheld. Galasso and Schankerman (2010) formalize this intuition, showing in the context of a bargaining game that fragmentation reduces the negotiation value of a patent that is potentially infringed and reduces the time to settlement in a patent dispute. In short, under this second view, the expected costs of infringement are highest when patents are highly concentrated, and the formation of a patent commons will have a particularly strong effect in reducing entry costs. We will allow the data to inform which of these alternative views shapes the implications of the patent commons, and so indirectly contribute to the ongoing debate on fragmentation of patent holdings and litigation risk.

3. Research Setting

3.1 The Commons

We focus on one major patent commons—the Open Source Development Labs’ Patent Commons project (commonly referred to as “The Commons”). In January 2005, IBM pledged access to 500 software patents to “any individual, community, or company working on or using software that meets the Open Source Initiative (OSI) definition of open source software now or in the future.” Subsequent to IBM’s action, several other incumbents pledged an additional 29 patents to The Commons.¹² “Pledge” in this context means that “patent holders agree they will not, under certain terms and conditions, assert patent rights against third parties who are engaging in activities that might otherwise give rise to a claim

¹² Example companies include Computer Associates International Inc. and Open Invention Network, LLC.

of patent infringement.”¹³ IBM announced in its press release that it believed this was the largest patent pledge of any kind. All pledged patents are explicitly listed on an online public database, and users of the technologies embedded in the patents are not required to sign any formal agreement with The Commons.

While we found a series of recent patent pledging events based on a search of major news outlets (see Appendix A for a detailed summary of these events), our choice of The Commons as the focus of our analysis is guided by several factors. First, The Commons is economically important in the sense that it comprises a large collection of patents. Second, The Commons covers multiple software technology markets, allowing us to use variation over time within software markets for identification. Third, The Commons was introduced in 2005, allowing sufficient time to observe changes in entry behavior after its introduction. Fourth, The Commons specifies contributed patents at a very detailed level, listing each of their patent numbers. We could find no other patent donation that similarly met all of these criteria. Nevertheless, in our empirical analysis we control for the effects of other commons-like institutions.

3.2 IBM historical support for open source software

The establishment of The Commons is consistent with IBM’s strategy since the late 1990s of supporting OSS (Campbell-Kelly and Garcia-Swartz 2009, Capeck et al. 2005, Samuelson 2006). IBM initially announced its commitment to Linux in 1999, and in 2001 announced that it would invest \$1 billion over the following three years to make Linux more suitable for enterprise applications (Campbell-Kelly and Garcia-Swartz 2009). In keeping with this commitment, IBM has made all of its hardware platforms compatible with Linux, released Linux versions of its software products, and developed Linux-focused service capabilities. IBM explicitly supports OSS to promote open standards in areas that are complementary to its profitable businesses (Capek et al. 2005). Over time, it has focused less on operating systems that might compete with open source alternatives, and focused more on developing and marketing middleware or application software. IBM’s business model now focuses on selling high end hardware, proprietary software running on top of Linux, and systems integration and other customized services to enterprise customers (Samuelson 2006).

IBM’s contributions to create and license OSS throughout our sample period may cast doubt about the real effects The Commons has on innovation by other OSS firms. As a practical matter, an OSS contributor implicitly grants a license to any patents that author holds that read on the OSS that has been developed; otherwise any users of that OSS may become inadvertent infringers of the patent. However, while some OSS licenses, such as the General Public License v3 (GPL v3), include an explicit grant of a patent license, many do not (Capek et al. 2005). Further, IBM may own IPRs that do not directly cover

¹³ For more details, see http://www.patent-commons.org/resources/about_commitments.php.

contributed technology but may be closely related to it. This creates legal uncertainties for both producers and users of OSS.

In summary, our analysis of IBM's involvement with OSS suggests that its support for OSS starts before and continues throughout our sample period (1999 to 2009, as explained below). However, the establishment of The Commons represents a significant discontinuity in terms of IBM's *explicit* legal support for OSS. The only similar event that occurs during our sample period was IBM Senior Vice President Nick Donofrio's announcement in August 4, 2004 during a keynote at the LinuxWorld conference that IBM would not assert its patents against the Linux kernel (Scannell 2004).¹⁴ We will examine the effect of this 2004 announcement and its relationship with the establishment of The Commons in 2005 when we study the robustness of our empirical results.

4. Data

4.1. Sample

Our sample consists of 2,054 US software firms from the 2004 and 2010 editions of the CorpTech Directory of Technology Firms¹⁵ (denoted as CorpTech 2004/2010 hereafter) that primarily operate in the prepackaged software industry. We combine this sample with data from the National Establishment Time Series (NETS) Database, which includes 100,000 US-based firms with primary SIC 7372. Our use of two data sources reflects constraints with each. The CorpTech data have detailed information on the product markets of firms, but have little time variation, while the NETS data have limited product market information but vary over time.

As noted above, the focus of our study is on start-up firms. As a result, we restrict our sample to firms that were founded after 1990 and that have fewer than 1000 employees and less than \$500 million in annual sales.¹⁶ Our sample period is from 1999 to 2009, with 6 years before the establishment of The Commons and 5 years after. We believe this time window is sufficiently long to capture the impact of The Commons on OSS entry.

4.2. Identifying software markets and the matching patent classes

We use the product code classification system embedded in the Gale database "PROMT" (Fosfuri et al. 2008) as our primary source to define software markets.¹⁷ Because of certain drawbacks of only relying on the PROMT classifications (we describe these in further detail in Appendix B), we further

¹⁴ It is believed that this announcement was made in response to a finding by the Open Source Risk Management organization that several large companies, including IBM, held patents that might affect the Linux kernel (Scannell 2004).

¹⁵ Our choice of 2010 CorpTech data reflects a constraint with the data—we have contacted CorpTech and there are no historical data from 2005 to 2009, the core years of our sample period. The combination of CorpTech 2004 and 2010 is to address potential survivor bias.

¹⁶ Our results are robust to the use of alternative thresholds for inclusion in our sample. For example, our results are robust to an alternative sample of start-ups that includes firms founded after 1990 that have fewer than 500 employees and less than \$500 million annual sales.

¹⁷ A few examples of PROMPT product codes are included in Table B-1 in Appendix B.

match PROMT’s software-related product categories with CorpTech’s “SOF” product classes¹⁸ to create a PROMT-CorpTech concordance so that each PROMT software product code is associated with a detailed set of keywords. The keywords for each product class are used to (i) manually assign PROMT product codes to PROMT news articles with missing codes and (ii) match software markets with the most relevant patent classes as described below.

An important part of our data construction involves matching product markets to patents. This allows us to identify both the cumulateness of innovation and the concentration of patent ownership in a software market. As is well-established in the literature, this type of matching is difficult (e.g., Griliches 1990, Silverman 1999). We follow Cockburn and MacGarvie (2006, 2011) and match software patents to CorpTech “SOF” product classes to create a patent-CorpTech concordance. Because our software markets are classified through PROMT categories, in order to create the final mapping between software markets and patent classes, we then combine the PROMT-CorpTech concordance and patent-CorpTech concordance to form the PROMT-patent concordance. The final concordance that we use in the empirical analysis consists of 33 software markets matched to 422 patent class-subclass combinations¹⁹ (see Appendix B for a detailed discussion of our data construction process).

5. Measures

5.1. Dependent variable: OSS entry

We measure OSS entry for software market j in year t as the number of events in which start-ups introduce their first OSS product into that market and year. We use a three-step procedure to identify OSS entry in a software market based on the press releases of the 2,054 firms in the PROMT database. First, following the work by Fosfuri et al. (2008) and Bessen and Hunt (2007), we searched for a set of keywords within PROMT articles to identify articles related to OSS. Appendix B includes the full set of keywords. Second, we *manually* read all search results that included words from the first step to identify new OSS product introductions. We identified an article as referring to an introduction of a new OSS product when the article indicated that either of the following took place: (i) the introduction of a new software product that offered one or more of its module(s)²⁰ under an open source license (we label such modules as *open source modules*); and (ii) the introduction of a new version of an existing software product with open source modules. Third, to identify OSS entry, we kept only the events in which the start-up introduced an open source module into a market in which it did not previously have OSS

¹⁸ There are more than 290 software product codes (denoted as SOF categories) defined by CorpTech Directory. Each firm in this directory is associated with a set of self-reported product codes selected from these 290 SOF categories. For further details on this classification system, see Cockburn and MacGarvie (2006).

¹⁹ Table B-3 in Appendix B lists examples of this final concordance between software markets and US patent class-subclass combinations.

²⁰ In software, a module is a part of a program. A software product is composed of one or more modules that are linked together but perform different functions (e.g., the calendar module available in the Microsoft Office's Outlook).

products. In total, we have 242 new OSS product entry events made by 85 start-up firms from 1999 to 2009.²¹ We aggregated these new OSS product entry events by software market and year. The data are structured as a balanced panel. Table 1 includes a brief description of measures and summary statistics for the main variables used in our empirical analysis.

Table 1: Summary Statistics

| Variable name | Measure (Market-year) | Obs. | Mean | Std. Dev. | Min | Max |
|---|---|------|--------|-----------|-------|--------|
| <i>Dependent variable</i> | | | | | | |
| OSS entry | The number of start-up entry events through releasing new OSS products related to market j in year t | 363 | .667 | 1.344 | 0 | 11 |
| <i>Independent variables and controls</i> | | | | | | |
| The Commons | Log of The Commons' claims-weighted patent count related to market j cumulated up to year t | 363 | 2.413 | 2.936 | 0 | 7.911 |
| Cumulativeness | Log of cumulativeness of innovation in market j up to year t | 363 | .808 | .613 | .095 | 3.454 |
| Concentration | Four-assignee citation concentration ratio in market j up to year t | 363 | 0.227 | 0.075 | 0.076 | 0.458 |
| Sales growth | The growth of market j 's sales from year $t-1$ to year t | 363 | 1.007 | 0.150 | 0.631 | 2.112 |
| Total patents | Log of total claims-weighted patent count related to market j cumulated up to year t | 363 | 10.817 | 1.232 | 6.870 | 13.486 |
| Patent quality | Log of average quality of patents related to market j cumulated up to year t | 363 | 2.832 | .402 | 1.839 | 4.051 |
| OIN patents | Log of Open Invention Network's claims-weighted patent count in market j cumulated up to year t | 363 | 1.125 | 2.053 | 0 | 6.690 |
| SSO patents | Log of standard-setting organizations' claims-weighted patent count in market j cumulated up to year t | 363 | 1.628 | 2.118 | 0 | 5.908 |
| OSS demand | Log of cumulated downloads at SourceForge for OSS applications related to market j up to year t | 363 | 14.340 | 5.721 | 0 | 22.365 |
| <i>Instrument variables</i> | | | | | | |
| IBM's opposed patents | IBM's patents opposed at the European Patent Office (EPO) related to market j cumulated up to year t | 363 | 5.804 | 5.666 | 0 | 28 |
| IBM's opposed patents X afteryear2003 | IBM's opposed patents * afteryear2003 dummy, where afteryear2003 dummy turns on for years 2003, 2004, ..., 2009 | 363 | 3.856 | 5.539 | 0 | 28 |

5.2. Independent variables

The Commons. This variable is equal to the log of claims-weighted patent counts in The Commons related to software market j in year t .²² We use the claims-weighted count to measure the size of The Commons for several reasons. First, since one potential effect of The Commons is to reduce invention and licensing costs for start-ups, using a claims-weighted measure may more precisely capture the scope of technologies donated by the incumbent. In particular, as noted in Cohen and Lemley (2001: p. 6), "the scope of a patent is defined by its claims, which set out each element of the invention".²³ Moreover, since patent claims reflect an inventor's effort to make the patent more resistant to invalidation challenges (Allison et al. 2004, Bessen 2008), the number of claims may serve as a proxy for patent value. In short,

²¹ This procedure implicitly assumes there is no OSS entry by firms prior to 1999. We believe this assumption is supported by empirical evidence. For example, SourceForge, a major repository of OSS, was started in November 1999.

²² We also use raw patent counts as a robustness check. The results are qualitatively similar to the claims-weighted measure and are reported in Table C-2 and Table C-3 in the Appendix.

²³ While we use patent claims as a proxy for patent scope, we by no means argue that the correlation is necessarily positive. For instance, Allison et al. (2004) suggests claims may be negatively correlated with patent scope.

by weighting by claims, we will be able to better capture variance in the value and scope of patents contributed to The Commons. We further take the logged value of this variable²⁴ to reduce skewness.

Cumulativeness. This variable refers to the cumulativeness of innovation within market j in year t . We use patents' backward citations, which provide information about "existing ideas used in the creation of new ideas" (Caballero and Jaffe 1993) and indicate "some form of cumulative technological impact" (Jaffe et al. 1998). Following Clarkson (2005), we measure it based on the average propensity for patents in market j and year t to backward cite patents within the same market j . This is roughly similar to the way economists have measured the cumulative nature of innovation at the firm level, e.g. using the extent to which firms self-cite their own patents (Hall et al. 2005). In our setting, we proceed as follows. If we sort the N patents within a software market j chronologically (with $m=1$ being the oldest patent and $m=N$ being the youngest), the cumulativeness for each patent n (i.e., the propensity for patent n to cite preceding patents within the same market) is calculated as $C_n = \frac{\sum_{m=1}^N x_{nm}}{n-1}$, where x_{nm} is a dummy variable equal to one if patent n back-cites patent m , and zero otherwise (with both patents belonging to the same market), $(n-1)$ is the total number of possible citations, and $n>1$, since C_1 is undefined. In other words, the cumulativeness of a focal patent in market j is based on the share of potential backward citations to patents belonging to the same market that are actually cited by the focal patent. The cumulativeness of innovation for software market j is then the average of all $N-1$ patents' cumulativeness:²⁵ $C_j = \frac{\sum_{n=2}^N \sum_{m=1}^N \frac{x_{nm}}{n-1}}{N-1}$. This measure varies over time based on the grant year of the market j patents under consideration.²⁶ Notice that the older patents in a market tend to have greater cumulativeness since the potential number of patents that can be cited is smaller. As a robustness check, we also used an alternative weighting scheme, one that provides relatively lower importance to the cumulativeness measure of older patents. As in Clarkson (2005), it is calculated as $C_j = \frac{\sum_{n=1}^N \sum_{m=1}^N x_{nm}}{N(N-1)/2}$. For both measures, we take the logged value to reduce skewness.

Concentration. This variable indicates the extent of concentration of patent ownership in a market. Following Noel and Schankerman (2006) and Cockburn and MacGarvie (2011), we use the four-assignee citation concentration ratio to measure the concentration of patent ownership in a software market. Backward citations indicate the extent to which a technological area has already been covered by prior art, so the share of backward citations owned by an assignee suggests the extent to which the assignee holds existing patented technologies and therefore the importance of negotiating with the assignee. To construct

²⁴ We add 1 to the variable when taking the log.

²⁵ The average only considers $N-1$ patents since C_1 is undefined.

²⁶ We acknowledge this measure also captures the extent to which knowledge is focused in a software market, which is slightly different from how cumulativeness of innovation is defined in this study. As a result, there may exist some nuance in interpreting the empirical results based on this measure.

this variable, we first calculate the number of citations made by patents in market j up to year t that are held by the cited assignee n (denoted as s_{njt}). Then we arranged s_{njt} in descending order. The total citations owned by the four firms that received the top four largest number of citations made by patents in market j in year t (i.e. the top four s_{njt} , where $n=1,2,3,4$) is $\sum_{n=1}^4 s_{njt}$. Thus, the four-assignee citation concentration ratio for market j in year t is calculated as $\frac{\sum_{n=1}^4 s_{njt}}{\text{total_citations}_{jt}}$, where $\text{total_citations}_{jt}$ is the total number of citations made by patents in market j up to year t .²⁷

5.3. Control variables

Sales growth. One important factor that may correlate with both the behavior of firms contributing to The Commons and OSS entry is the rate of market growth for software market j , which is proxied by the sales change from year $t-1$ to year t in market j . Because we do not have CorpTech data between 2005 and 2009, we use NETS data to measure this variable. Roughly 4,500 software firms in the NETS data are assigned to one of the eight-digit SIC categories (e.g., 73729901) that correspond to eight broad categories in the software industry. We compute the yearly sales change for each of the eight SIC categories and then map them to our 33 software markets to approximate the overall sales growth for each market for a given year.

Total patents. Although we are most interested in two of the most important features of patent thickets—the cumulativeness of innovation and the concentration of patent ownership, the total number of patents related to a market has also been used as a measure of the density of patent thickets (Cockburn and MacGarvie 2011). We add this variable as an additional control and measure it using the cumulative number of granted patents that are not included in The Commons but related to software market j . To be consistent with how we measure The Commons, we use the log of the claims-weighted patent count.

Patent quality. This variable is a control for the quality of patents in market j in year t . As has been noted elsewhere, higher quality patents suggest superior technological capabilities possessed by existing incumbents in the market, which leaves less room for start-ups to innovate further. This variable is equal to the log value of the cumulative stock of citations received by the patents in market j (adjusted for truncations) divided by total number patents in j up to year t .

Open Invention Network (OIN) patents. At the end of our sample period, another institution similar to patent commons—OIN—was established. Similar to The Commons, OIN offers contractually royalty-free usage of its patents to OSS participants as long as users promise not to file suit against software associated with the Linux System.²⁸ We do not focus on the implications of the introduction of

²⁷ We also use an eight-assignee citation concentration ratio as a robustness check. The results are qualitatively similar to this four-assignee citation concentration ratio measure.

²⁸ For the detailed definition of the Linux system, see http://www.openinventionnetwork.com/pat_linuxdef.php.

this institution in our main analysis as it was introduced too late in our sample period to have a measureable effect.²⁹ However, we include it as a control. We measure this variable as the log of the claims-weighted patent count of OIN patents related to software market j cumulated up to year t .

Standard-setting organization (SSO) patents. As mentioned earlier, another important mechanism to address the anti-commons problem is SSOs. Such institutions promote coordination of innovation by providing a forum for collective decision-making among firms, facilitating the introduction of standards (Rysman and Simcoe 2008). If any patent is incorporated into the standards, the patent owner can gain significant power to control the diffusion of such standards and even deter market entry (Shapiro 2001, Rysman and Simcoe 2008). To prevent this blocking effect, most SSOs require patent holders contributing to the standard to license their patents on “Fair, Reasonable, and Non-Discriminatory (FRAND)” terms. Firms can even choose to license their patents on royalty-free terms. We control for the incidence of SSO patents that are licensed royalty-free because we expect that such patents might also have some effect on OSS entry. Therefore, we collect all patents disclosed under royalty-free licenses by the major eight SSOs (e.g., IEEE, ITU) from 1971 to 2008 (Rysman and Simcoe 2008)³⁰. We compute the claims-weighted patent count of the SSO patents that are distributed under royalty-free licenses and are related to software market j cumulated up to year t .

OSS demand. While *Sales growth* could potentially control for the market growth for software market j in year t in general, the market demand for OSS products may deviate from the demand for proprietary software products. We capture the demand for OSS using the total OSS downloads at SourceForge.net (SourceForge), which is the largest repository for OSS projects over the world. We first match over 0.2 million OSS projects at SourceForge to the 33 software markets according to the projects’ categories. Then we measure this variable by computing the total number of downloads for all OSS applications related to software market j cumulated up to year t .

6. Empirical Strategies and Results

We first motivate our empirical analyses by characterizing the quality of patents in The Commons as well as their distribution across markets, and then document the characteristics of software markets before and after the formation of The Commons for markets where The Commons has few and many patents. Next, we establish a baseline relationship between OSS entry and the size of The Commons, and demonstrate how the marginal impact of The Commons varies with the cumulativeness of innovation or patent ownership concentration. We then show that our baseline results are robust to a variety of specifications and robustness checks.

²⁹ For the 130 patents contributed to OIN from year 2006 to 2009, 70 percent were contributed in 2008 and 2009.

³⁰ We are grateful to the generous offer of the SSO patent data set by Tim Simcoe and Christian Catalini. These data are available for download under a creative commons license at www.ssopatents.org.

6.1. Characteristics of software markets and patents in The Commons

Figure 4 below shows the distribution of patents in The Commons across 33 software markets. The markets with the greatest concentration of patents from The Commons are operating systems and utilities, disk/file management, and database software, followed by software development tools and system management software. Because 95% of patents in The Commons were contributed by IBM, we compared it with the distribution of IBM's patent holdings across market, and the two are quite similar.

We next examine the quality of patents in The Commons relative to comparison groups. We first compare patents in The Commons to similar market patents. Following the matching method employed by Thompson and Fox-Kean (2005), we pair each patent in The Commons with a randomly selected control patent that matches the primary classification of The Commons' patent at the subclass level and that was applied for in the same year or within a one year window. As shown in Table 2a, there is no statistically significant difference between the two groups in terms of forward citations or backward citations. The patents in The Commons have lower number of claims than other market patents, suggesting The Commons may be less resistant to invalidation challenges (Bessen 2008).

We next compare the quality of IBM's donated patents to other patents in its portfolio using the same procedure described above. As shown in Table 2b, there is no significant difference in forward citations between the two groups. Similar to the comparison with similar market patents, the patents in The Commons have a slightly lower number of claims than IBM's other patents. Moreover, the patents in The Commons show a lower number of backward citations, which may suggest these patents are less derivative than comparable IBM patents (Lanjouw and Schankerman 2004). Nevertheless, the lower number of backward citations made by The Commons' patents could also imply they are more subject to the challenges by prior art they possibly missed (Bessen 2008).

In general, these comparisons suggest that while patents in The Commons may be less resistant to invalidation challenges, they have similar quality and may be somewhat less derivative.

We also examine the characteristics of markets where the number of claims-weighted patents in The Commons is small (below the 25th percentile) and where it is large (above the 75th percentile), before and after the Commons was introduced. As in Table 3, in markets where The Commons patents are well represented, OSS entry is greater during both pre-Commons and post-Commons period. However, during the pre-Commons period, the sales growth is lower in markets where the incumbents were about to donate large number of patents to The Commons; it became greater in these markets after The Commons was established. In markets where The Commons patent are well represented, cumulativeness of innovation and patent quality are lower while patent ownership concentration and total number of market patents are greater. These differences persist before and after The Commons was introduced.

Figure 4: Distribution of patents in The Commons and in IBM's portfolio by market

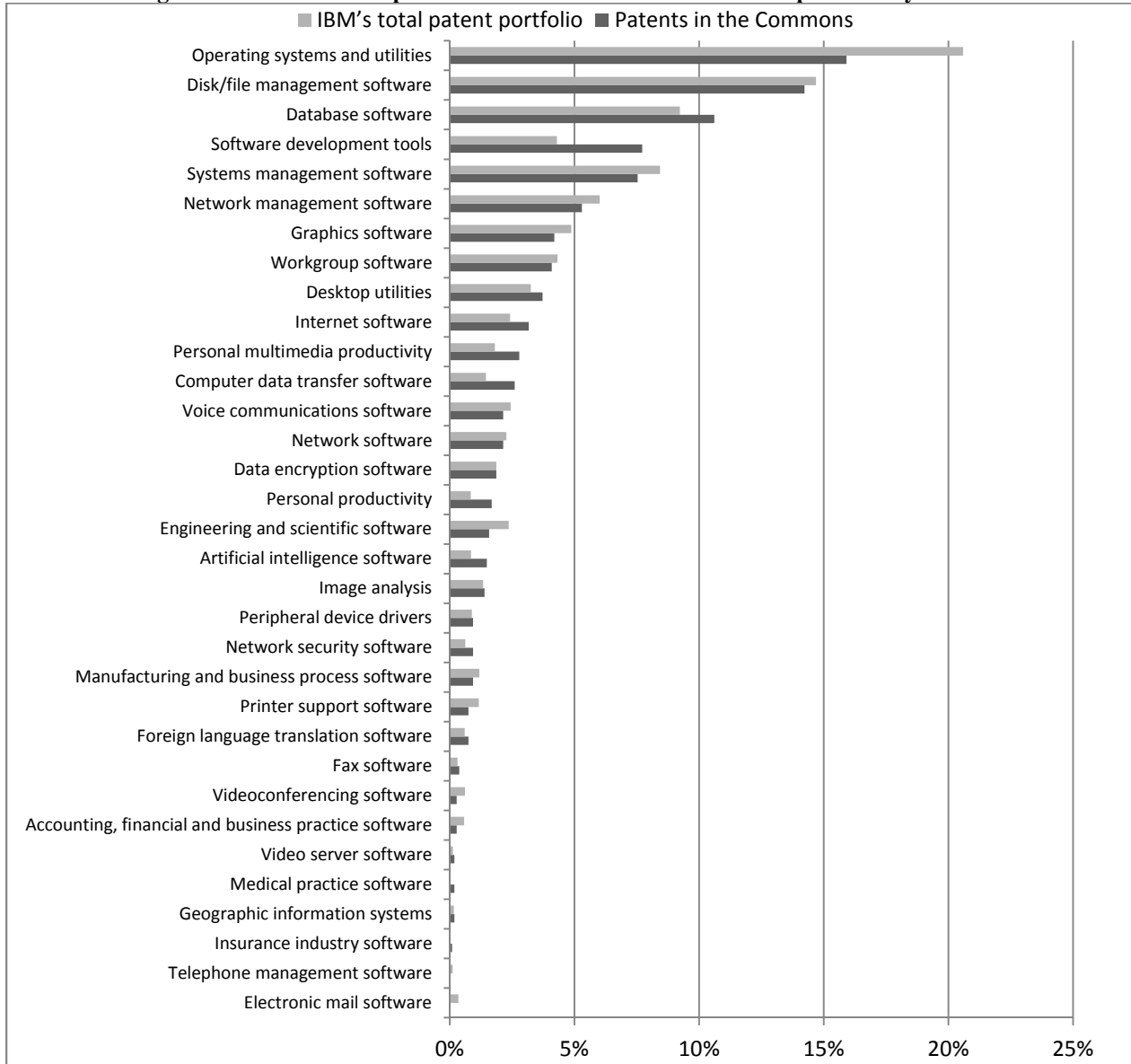


Table 2a: Patents in The Commons compared to other similar market patents

| | | Patents in The Commons | Other similar market patents | T-test |
|----------------------------------|-----------------|------------------------|------------------------------|-----------|
| | Obs. | 517 | 517 | |
| Forward citations as of Dec 2004 | Mean (Std.Err.) | 11.393 (.683) | 12.334 (.788) | -0.903 |
| Backward citations | Mean (Std.Err.) | 9.718 (.314) | 9.739 (.394) | -0.042 |
| Claims | Mean (Std.Err.) | 17.321 (.496) | 19.400 (.686) | -2.458*** |

Note: 1) For each of the patents in The Commons, we paired it by randomly selecting a control patent from all market patents that matched the primary classification of The Commons patent at the subclass level and that were applied in the same year or within one-year window of The Commons patent. 2) To eliminate any effect of contributing to The Commons on the increase in forward citation, when comparing the forward citations, we only count the total forward citations received by December 2004, as The Commons was formed in January 2005. Because 57% of the patents in Commons are granted before 1998 and 97% of the patents are granted before 2002, we believe the truncation on forward citations is not a major issue in this setting. 3) ***: significant at 1%.

Table 2b: IBM's Patents in The Commons compared to IBM's other similar patents

| | | IBM's Patents in The Commons | IBM's other similar patents | T-test |
|----------------------------------|-----------------|------------------------------|-----------------------------|-----------|
| | Obs. | 407 | 407 | |
| Forward citations as of Dec 2004 | Mean (Std.Err.) | 11.044 (.784) | 10.341 (.702) | 0.668 |
| Backward citations | Mean (Std.Err.) | 9.683 (.358) | 11.545 (.509) | -2.992*** |
| Claims | Mean (Std.Err.) | 16.813 (.501) | 18.646 (.612) | -2.319** |

Note: 1) For each of IBM's patents in The Commons, we paired it by randomly selecting a control patent from all IBM's patents that matched the primary classification of the focal patent at the subclass level and that were applied within one-year window of The Commons patent. 2) To eliminate any effect of contributing to The Commons on the increase in forward citation, when comparing the forward citations, we only count the total forward citations received by December 2004, as The Commons was formed in January 2005. Because all IBM's patents in Commons are granted in year 1993, 1997, or 2001, we believe the truncation on forward citations is not a major issue in this setting. 3) ***: significant at 1%, **: significant at 5%.

Table 3: Market characteristics in pre-Commons and post-Commons periods

| | Pre-Commons (1999-2004) | | | Post-Commons (2005-2006) | | |
|----------------|------------------------------------|------------------------------------|-----------|------------------------------------|------------------------------------|------------|
| | Markets Where The Commons is small | Markets Where The Commons is large | T-test | Markets Where The Commons is small | Markets Where The Commons is large | T-test |
| OSS entry | .278 (.100) | 1.389 (.273) | 3.823*** | .155 (.055) | 1.844 (.330) | 5.045*** |
| Cumulativeness | 1.506 (.107) | .552 (.045) | -8.188*** | 1.091 (.094) | .326 (.029) | - 7.745*** |
| Concentration | .213 (.011) | .295 (.007) | 6.729*** | .182 (.012) | .256 (.007) | 5.557*** |
| Sales growth | 1.051 (.013) | .985 (.021) | -2.616*** | .978 (.020) | 1.041 (.028) | 1.801** |
| Total patents | 9.392 (.154) | 11.262 (.118) | 9.667*** | 10.286 (.136) | 12.207 (.115) | 10.806*** |
| Patent quality | 3.253 (.054) | 3.121 (.031) | -2.116** | 2.665 (.045) | 2.553 (.026) | -2.156** |

Note: 1) Markets where The Commons is small (large) are those where the number of claims-weighted patents in The Commons are below the 25th percentile (above the 75th percentile). 2) The definitions of these market characteristics are shown in Table 1. 3) ***: significant at 1%, **: significant at 5%.

6.2. Baseline results and robustness checks

We first seek to establish whether increases in the size of The Commons are associated with increasing entry in related markets. Our estimation framework is motivated by recent research that has studied how patent thickets influence market entry in the software industry (e.g., Cockburn and MacGarvie 2011). We model OSS entry using count data models with conditional fixed effects. Suppose the number of OSS entry events in software market j in year t (denoted as Y_{jt}) follows a Poisson process with parameter λ_{jt} taking the form $\lambda_{jt} = \exp(X_{jt}'\beta)$. Also suppose α_j is a market-specific and time-constant variable that incorporates unobserved heterogeneity across markets. Thus, $E(Y_{jt} | X_{jt}, \alpha_j) = \lambda_{jt} = \alpha_j \exp(X_{jt}'\beta)$, where $X_{jt}'\beta = \beta_1 The Commons_{jt} + \gamma_1 PatentThicket_{jt-1} + \gamma_2 SalesGrowth_{jt} + \gamma_3 MarketPatents_{jt-1}$

$$+ \gamma_4 OtherFreePatents_{jt} + \gamma_5 OSSdemand_{jt} + \tau_t \quad (1)$$

The vector $PatentThicket_{jt-1}$ includes the two patent thicket variables $cumulativeness_{jt-1}$ and $concentration_{jt-1}$; the vector $MarketPatents_{jt-1}$ includes the $Total patents_{jt-1}$ and $Patent quality_{jt-1}$. The two vectors are lagged by one year to allow for any lagged effects on OSS entry. The vector $OtherFreePatents_{jt}$ represents the patents from OIN and SSOs— $OIN patents_{jt}$ and $SSO patents_{jt}$. τ_t includes 10 year dummies to control for time-varying factors that may influence OSS entry. The model is then estimated using maximum likelihood with robust standard errors clustered at the market level. We are interested in the estimate for β_1 which, if positive, suggests The Commons have some positive impact on OSS entry. Our identification assumption in this section is that there are no unobserved market-level factors that are correlated both with entry and with the size of The Commons. This assumption will be

violated and thus the estimate will be biased if, for example, the contribution decision by IBM and other contributors is driven by market growth that is not adequately captured by our control variables.³¹ We will examine the implications of weakening this assumption in a variety of ways in the analyses below.

Columns (1) and (2) in Table 4 report the estimation results for specification (1). Column (1) shows the specification with baseline controls including $SalesGrowth_{jt}$, $PatentThicket_{jt-1}$, and $MarketPatents_{jt-1}$ as well as with market and year fixed effects. We include $SalesGrowth_{jt}$ in all specifications because it is an important determinant of start-up entry.³² We further control for market-level patent holdings through the controls $PatentThicket_{jt-1}$ and $MarketPatents_{jt-1}$. The coefficient in column (1) suggests that a 10% increase in The Commons' patent claims related to a software market is associated with a 1.3% increase in OSS entry in that market. Results are robust when we add other controls such as $OtherFreePatents_{jt}$ and $OSSdemand_{jt}$, as shown in column (2) in Table 4.

We next investigate whether the impact of The Commons is greater in markets with high cumulateness of innovation and the concentration of patent ownership in a market. Our key identification assumption here is somewhat weaker—that growth in claims-weighted patents in The Commons are not correlated with unobservable factors that influence entry that are differentially trending in markets with high cumulateness and concentration. We begin by examining whether the impact of The Commons is higher when the cumulateness of innovation in a market is high. The specification for $X_{jt}'\beta$ becomes:

$$X_{jt}'\beta = \beta_1 The\ Commons_{jt} + \beta_2 The\ Commons_{jt} * cumulateness_{jt-1} + \gamma_1 PatentThicket_{jt-1} + \gamma_2 SalesGrowth_{jt} + \gamma_3 MarketPatents_{jt-1} + \gamma_4 OtherFreePatents_{jt} + \gamma_5 OSSdemand_{jt} + \tau_t. \quad (2)$$

As shown in columns (3) and (4) in Table 4, a 10% increase in the size of The Commons is associated with a 2.8%-3% increase in OSS entry, with the effect computed at the average level of cumulateness of innovation. Further, while the marginal effect of The Commons is insignificant when evaluated at the 10th percentile of the cumulateness, the effects are statistically and economically significant when evaluated at the 90th percentile. Specifically, a 10% increase in the size of The Commons is associated with a 5.2%-5.4% increase in OSS entry when cumulateness of innovation is at its 90th percentile. A test for the difference of the two marginal effects (at the 10th and 90th percentiles) is statistically significant. As described earlier, we also constructed alternative measures of the cumulateness of innovation and our results are robust to these changes.³³ The results suggest that the impact of The Commons is greater when the cumulateness of innovation is high.

³¹ It is worth noting that there could be multiple unobserved market characteristics that may bias the estimates upward or downward. For example, if IBM and other incumbents are interested in contributing patents to markets where the OSS growth is the fastest, the estimated coefficient is biased upward. On the other hand, if IBM and other incumbents intend to promote the development of OSS, it may have to focus on markets where the OSS growth is slow, leading to a downward estimate.

³² We have experimented with other controls for market growth such as the number of incumbents and sales volume. Regressions using these other controls yield qualitatively similar results for the main parameters of interest.

³³ The regression results based on the robust measure of cumulateness are available upon request.

To explore how the effect of The Commons on patent ownership concentration, the specification for $X_{jt}'\beta$ is: $X_{jt}'\beta = \beta_1 The\ Commons_{jt} + \beta_2 The\ Commons_{jt} * concentration_{jt-1} + \gamma_1 PatentThicket_{jt-1} + \gamma_2 SalesGrowth_{jt} + \gamma_3 MarketPatents_{jt-1} + \gamma_4 OtherFreePatents_{jt} + \gamma_5 OSSdemand_{jt} + \tau_t$. (3)

The empirical results for this specification are reported in columns (5) and (6) in Table 4. While the marginal effect of The Commons is insignificant when concentration is at its 10th percentile or mean value, a 10% increase in the size of The Commons is associated with a statistically significant 1.1%-1.2% increase in OSS entry when concentration is at its 90th percentile. The test for the difference of marginal effects of The Commons between concentration evaluated at the 10th percentile and the 90th percentile is statistically significant at the 10% level.

Table 4: Baseline results, conditional fixed-effect Poisson regression

| Dependent variable: OSS entry | Specification testing direct impact of The Commons only | | Add interaction with cumulativeness of innovation | | Add interaction with concentration of patent ownership | | Add interaction with cumulativeness of innovation and with concentration of patent ownership | |
|----------------------------------|---|-------------------|---|---------------------|--|--------------------|--|---------------------|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| The Commons | .127* (.083) | .145* (.087) | .033 (.119) | .052 (.124) | -.079 (.114) | -.098 (.135) | -.080 (.120) | -.071 (.164) |
| The Commons * Cumulativeness | | | .330*** (.096) | .331*** (.098) | | | .308*** (.095) | .306*** (.100) |
| The Commons * Concentration | | | | | .566* (.322) | .660* (.369) | .309 (.270) | .336 (.377) |
| Cumulativeness | .965 (.977) | 1.010 (1.156) | 3.194*** (.858) | 3.303*** (.983) | 1.063 (.886) | 1.030 (1.024) | 3.116*** (.854) | 3.142*** (.990) |
| Concentration | -6.961 (5.273) | -7.625 (7.219) | -11.238** (4.812) | -12.196* (7.097) | -11.293* (6.552) | -10.709 (7.459) | -13.251** (5.539) | -13.375* (6.878) |
| Sales growth | -.412 (.338) | -.400 (.366) | -.571* (.379) | -.567 (.406) | -.454 (.354) | -.470 (.381) | -.585* (.384) | -.590 (.415) |
| Total patents | -.707 (1.408) | -.713 (1.505) | .797 (1.430) | .852 (1.465) | -1.096 (1.443) | -1.157 (1.560) | .510 (1.439) | .527 (1.519) |
| Patent quality | -1.435 (2.453) | -1.599 (2.387) | -1.224 (2.401) | -1.361 (2.335) | -1.529 (2.523) | -1.684 (2.422) | -1.284 (2.445) | -1.401 (2.350) |
| OIN patents | | -.066 (.071) | | -.069 (.065) | | -.076 (.070) | | -.073 (.066) |
| SSO patents | | -.010 (.078) | | -.016 (.065) | | .029 (.083) | | .005 (.078) |
| OSS demand | | .062 (.095) | | .055 (.083) | | .062 (.100) | | .056 (.086) |
| Log pseudolikelihood | -227.909 | -227.351 | -223.884 | -223.327 | -226.961 | -226.270 | -223.630 | -223.082 |
| <i>Marginal Effects</i> | | | | | | | | |
| The Commons (average) | .127* (.083) | .145* (.087) | .278** (.115) | .297** (.120) | .046 (.072) | .048 (.081) | .217** (.085) | .230** (.114) |
| The Commons (cumulativeness=10%) | | | .095 (.114) | .114 (.119) | | | .047 (.091) | .061 (.110) |
| The Commons (cumulativeness=90%) | | | .519*** (.149) | .540*** (.154) | | | .442*** (.119) | .454*** (.153) |
| The Commons (concentration =10%) | | | | | -.005 (.084) | -.011 (.098) | .190** (.091) | .200* (.134) |
| The Commons (concentration =90%) | | | | | .104* (.071) | .115* (.074) | .249*** (.086) | .265*** (.100) |

Notes: 1) All regressions include year dummies and market fixed effects. Number of observations=286. This number of observations is lower than 363 because of the use of conditional fixed effects Poisson models, which drops markets without OSS entry over the entire sample period. 2) Robust standard errors, clustered by market, are in parentheses. 3) * significant at 10%, ** significant at 5%, *** significant at 1%.

In columns (7) and (8) of Table 4 we present the results of a specification including the interactions with both cumulateness and concentration. In these models, a 10% increase in the size of The Commons is associated with a 4.4%-4.5% increase in OSS entry when the cumulateness of innovation is at its 90th percentile; the marginal effect of The Commons is significantly different at the 1% level when evaluated at high versus low level of cumulateness. In this specification, while there is no statistically significant difference between the marginal effects evaluated at the 10th and 90th percentiles of concentration, the sign for the interaction between The Commons and patent ownership concentration remains positive across specifications. We suspect this non-significance is likely caused by the multicollinearity between the two interactions.³⁴

We implemented a series of additional analyses to study the robustness of our results. We estimated all of our models using OLS models with market fixed effects. We substituted our claims-weighted count of number of Commons patents in the market with a raw patent count. Because of the decline in the rate of new OSS product entry after 2006 (as shown in Figure 5 below), we re-estimated the baseline model using a sample endpoint of 2007 and 2008. As detailed in Appendix C Tables C-1 through C-4, our results are robust to all of these changes.

Because our measure of patent ownership concentration could be confounded with the concentration of market structure, we constructed a measure of market structure concentration using the share of top 4 incumbents' sales in each market in each year. We found that adding this measure as an additional control did not affect our finding that The Commons had a stronger positive effect on entry in markets with high patent concentration. Further, we interacted the new market structure control with The Commons variable, in regressions both including and excluding our interaction of patent ownership concentration with The Commons. In both Poisson regression (in Table C-5) and linear models (in Table C-6) we find that changes in market structure concentration have no significant impact on the marginal effect of increases in the size of The Commons. However, even when controlling for the effects of market structure, changes in The Commons have a stronger effect on entry when patent concentration is high. This set of tests supports the view that it is the concentration of IPR holdings, not concentration of market structure, that influences the marginal effect of The Commons on OSS entry.

6.3. Examining the potential implications of omitted variables

This section provides the results of a variety of additional tests we run to examine how time-varying omitted variables might influence our results. We first control for time-varying market-specific trends using a model that includes a linear time trend interacted with the market fixed effects. The results are presented in Table 5.

³⁴ In the pooled sample, the simple correlation coefficient between the two interaction terms is 0.66.

In column 1 we show the results of a model with market-specific time trends but excluding the interactions with cumulateness and concentration. The results are qualitatively similar to our baseline, though the size of the marginal effect is somewhat smaller. We next interact the size of The Commons with cumulateness and concentration. Consistent with the baseline results, while the marginal effect of The Commons is insignificant when either cumulateness of innovation or concentration is at its 10th percentile, a 10% increase in the size of The Commons is associated with a 2%-4% increase in OSS entry when cumulateness is at its 90th percentile and is associated with a 1.1%-1.5% increase in OSS entry when concentration is at its 90th percentile, as shown by columns (2) through (4). The test of the difference between high and low cumulateness and the test of the difference between high and low concentration are also largely consistent with the baseline results.

Table 5: Robustness test using market-specific time trends, conditional fixed-effect Poisson regression

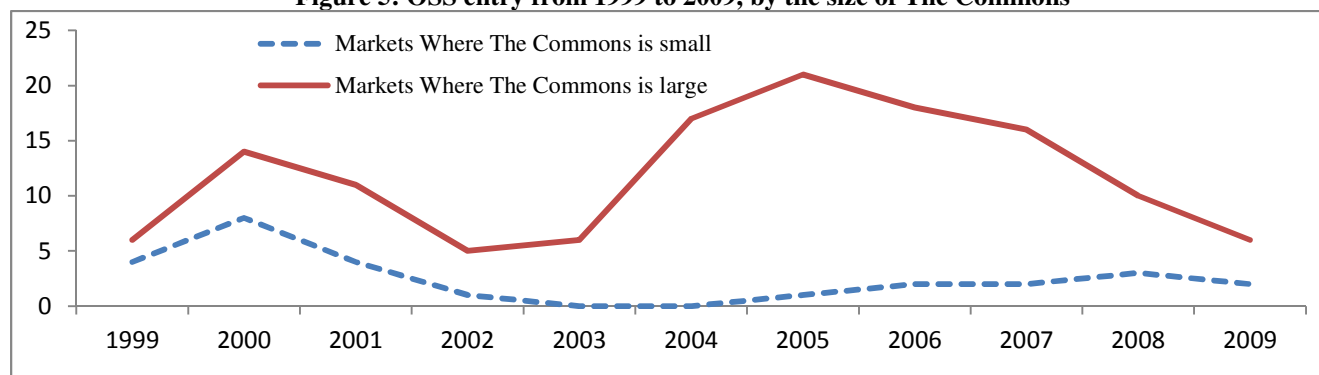
| Dependent variable: OSS entry | Specification testing direct impact of The Commons only | Add interaction with cumulateness of innovation | Add interaction with concentration of patent ownership | Add interaction with cumulateness of innovation and with concentration of patent ownership |
|----------------------------------|---|---|--|--|
| | (1) | (2) | (3) | (4) |
| The Commons | .093** (.041) | -.048 (.079) | -.235** (.102) | -.267 (.183) |
| The Commons * Cumulateness | | .298** (.118) | | .172* (.112) |
| The Commons * Concentration | | | 1.072*** (.389) | .909 (.663) |
| Cumulateness | | .612 (1.380) | | -.273 (1.323) |
| Concentration | | | 17.016** (8.210) | 17.734 (7.912) |
| Log pseudolikelihood | -231.251 | -229.011 | -223.827 | -223.018 |
| <i>Marginal Effects</i> | | | | |
| The Commons (average) | .093** (.041) | .174*** (.045) | .002 (.044) | .062 (.057) |
| The Commons (cumulateness=10%) | | .009 (.061) | | -.033 (.067) |
| The Commons (cumulateness=90%) | | .391*** (.114) | | .188* (.115) |
| The Commons (concentration =10%) | | | -.093 (.060) | -.019 (.105) |
| The Commons (concentration =90%) | | | .112** (.055) | .155*** (.052) |

Notes: 1) All regressions include market-specific time trends and market fixed effects. Number of observations=286. 2) Robust standard errors, clustered by market, are in parentheses. 3) “Market-specific time trend” is measured by a linear time trend times 33 market dummies. 4) Given our sample size of only 286 observations, it is difficult to incorporate any other controls besides these 33 market-specific time trends. 5) * significant at 10%, ** significant at 5%, *** significant at 1%.

Next we conduct a series of falsification exercises to see whether our results are consistent with a causal interpretation. We first examine whether the effects of the Commons exhibit the “right” timing. In particular, we should not observe a benefit from the intensity of contributions to The Commons prior to when it was initiated. Figure 5 shows the time series of OSS entry for markets with small and large size of The Commons. The Figure shows little variance over time in entry patterns in market segments with few (claims-weighted) Commons patents. The Figure shows that markets where Commons contributions are

large see little benefits to entry between 1999-2003, however there is a significant increase in entry in 2004, one year prior to the introduction of The Commons.

Figure 5: OSS entry from 1999 to 2009, by the size of The Commons



Notes: Markets where The Commons is small (large) are those where the number of claims-weighted patents in The Commons are below the 25th percentile (above the 75th percentile).

We believe this result reflects the effects of other actions taken by IBM in support of OSS shortly before the introduction of The Commons. In particular, as mentioned earlier, one significant event in the OSS community during this time period was the March 2003 lawsuit filed against IBM for alleged incorporation of SCO’s proprietary UNIX code into the open-source Linux operating system. This lawsuit, together with a well-publicized report in 2004 that suggested that Linux was potentially infringing on several hundred patents³⁵, initiated a series of actions by IBM that eventually lead to the introduction of The Commons and other actions taken to provide additional legal protection for OSS firms. This included the initiation by IBM of a legal defense fund for Linux in January 2004 and IBM Senior Vice President Nick Donofrio’s announcement that IBM would not assert its patents against the Linux kernel in August 2004. The appearance of these other events may cause us to identify an effect for The Commons before it actually appears. Further, Figure 4 suggests IBM’s donations to The Commons were highly correlated with its technological holdings as proxied by its patent portfolio. As a result, the pattern observed in Figure 5 may capture the effects of current and expected future activities by IBM in support of OSS that are correlated with IBM’s technological presence. This result would be consistent with our discussion in section 2 that a patent commons may provide a signal of the donator’s commitment to providing support to the OSS community. Table 6 shows the results of a series of specifications that we use to explore how these other events might influence our results.

We first construct a variable, denoted as “*False Commons*”, by interacting The Commons with a time dummy that turns on for years 2003 and 2004. Column (1) of Table 6 shows that if we add this variable to our baseline specification the coefficient for *False Commons* is significantly positive, which is

³⁵ See Alexy and Reitzig (2013) for detailed discussion about this event.

consistent with the descriptive evidence summarized in Figure 5. Column (2) shows that *False Commons* does not differentially affect OSS entry in markets with high cumulateness and concentration, however.

To explore whether the results in columns (1) and (2) reflect other activities by IBM in support of open source, we add to regression model in column (1) the cumulative number of claims-weighted *IBM patents* in the market that are not in The Commons and its interaction with a dummy that turns on for years 2003 and 2004 (denoted as *IBM patents X 2003_2004*) As shown in column (3), the coefficient on *IBM patents* is large and statistically significant, possibly picking up changes in OSS market opportunities over time across segments. Moreover, *IBM patents X 2003_2004* is significantly and positively correlated with OSS entry, reflecting how IBM's other actions during this time of legal uncertainty affected OSS entry. Despite the addition of these variables, the effects of *The Commons* remain statistically significant and its interactions with cumulateness and concentration (cf. column 4) are qualitatively similar.

Table 6: Falsification test by adding “False Commons”, conditional fixed-effect Poisson regression

| Dependent variable: OSS entry | (1) | (2) | (3) | (4) | (5) | (6) |
|----------------------------------|---------------------|----------------------|----------------------|----------------------|----------------------|-----------------------|
| The Commons | .373*** (.117) | .168 (.178) | .383*** (.140) | .093 (.164) | .426*** (.134) | .209 (.172) |
| False Commons | .463*** (.137) | .623** (.282) | | | .260 (.200) | .506 (.324) |
| IBM patents | | | .908*** (.249) | .867*** (.203) | .909*** (.257) | .855*** (.202) |
| IBM patents X 2003_2004 | | | .565*** (.160) | .504*** (.184) | .308 (.251) | .297 (.404) |
| The Commons X Cumulateness | | .207 (.132) | | .076 (.121) | | .120 (.136) |
| False Commons X Cumulateness | | .026 (.136) | | | | .077 (.208) |
| The Commons X Concentration | | .302 (.470) | | .646 (.405) | | .423 (.465) |
| False Commons X Concentration | | -.531 (.670) | | | | -.638 (.703) |
| Cumulateness | -.272 (1.174) | 1.371 (1.402) | -.145 (1.113) | .497 (1.217) | -.413 (1.061) | .546 (1.420) |
| Concentration | -12.659* (6.569) | -17.654** (7.055) | -14.613** (7.245) | -17.900** (6.933) | -15.932** (6.949) | -20.170*** (7.029) |
| Log pseudolikelihood | -222.405 | -219.586 | -219.448 | -218.066 | -218.754 | -216.825 |
| <i>Marginal effects</i> | | | | | | |
| The Commons (average) | .373*** (.117) | .389** (.119) | .383*** (.140) | .292*** (.121) | .426*** (.134) | .392*** (.124) |
| The Commons (cumulateness=10%) | | .275** (.127) | | .250** (.127) | | .326** (.128) |
| The Commons (cumulateness=90%) | | .540*** (.168) | | .347** (.162) | | .480*** (.177) |
| The Commons (concentration =10%) | | .362** (.140) | | .234* (.138) | | .355** (.144) |
| The Commons (concentration =90%) | | .420*** (.111) | | .358*** (.113) | | .435*** (.116) |

Notes: 1) All regressions include year dummies and market fixed effects. Number of observations=286. 2) All regressions include the full set of controls. The complete results including estimates for the controls are reported in Appendix Table C-7. 3) Robust standard errors, clustered by market, are in parentheses. 4) * significant at 10%, ** significant at 5%, *** significant. 5) *False Commons* is measured by interacting the patent density in The Commons across markets with a time dummy that turns on for year 2003 and 2004. 6) *2003_2004* is a dummy that turns on for years 2003 and 2004. 7) *IBM patents* is measured by log of claims-weighted patent count of all IBM patents (excluding IBM patents contributed to The Commons) related to market j cumulated up to year t.

In columns (5) and (6) we include *False Commons*, *IBM patents*, and *IBM patents X 2003_2004*. After controlling for the effect of IBM's IPR commitments and its actions in 2003 and 2004, *False Commons* becomes insignificant while *The Commons* (whose effect "turns on" in 2005) remains significantly positive. These findings corroborate our hypothesized effect, suggesting that IBM's pledge to The Commons go beyond these earlier OSS support actions and announcements, by identifying a specific set of complementary technologies that may be valuable to Linux and existing OSS projects.

The next set of falsification exercises examines the effects of The Commons on type of market entry. Because the introduction of The Commons will only benefit producers of OSS, we should not observe a positive effect of The Commons on entry of proprietary software products. Again using PROMT articles, we identified entry with proprietary products by start-ups into 29 software markets from year 2002 to year 2009.³⁶ Unlike our analysis using OSS entry, for this entry variable we measured entry using only PROMT product codes. That is, because of the very large number of potential entrants, we were unable to supplement our variable construction through manual reading of product articles as we did for OSS. This is a significant limitation, as 45% of product introduction events are missing PROMT product codes. Still, while this variable is measured with error, we are able to capture significant variance in proprietary product entry in our sample. The mean number of proprietary entry events per software market per year is 10.276 for proprietary software, compared with a mean number of 0.746³⁷ for OSS entry identified through both PROMT code and manual reading and a mean number of 0.134 for OSS entry identified using PROMT codes only. This comparison suggests that we are able to capture significant proprietary product entry in our data.

The results of this analysis, reported in Table 7, suggest that The Commons has no significant effect on proprietary entry; further, the interaction of The Commons with cumulativeness or concentration plays little role in proprietary entry as well.

Throughout this paper we have used cross-sectional variance in the intensity of contribution to The Commons to identify its effects across OSS markets. Another identification approach is to assume that the benefits of The Commons are equal across markets (i.e., they do not vary with intensity of contribution) and to use differences in product entry across equivalent OSS and proprietary markets to identify the effects of The Commons. Specifically, suppose each software market category (e.g., operating

³⁶ We use the press releases of the 2,054 start-up firms from 2002 to 2009 in the PROMT database to identify new proprietary software product entry. For each start-up, we only use the firm's first product in a market to capture entry. This results in 2,384 proprietary product entry events. Because of data limitations, we have incomplete product code information for most press releases before 2002, forcing us to focus on years from 2002 to 2009 only. We also adjust for a change in the assignment of product codes during our sample. Specifically, between 2007 and 2009 application-related software products were systematically assigned to a higher product code level in the PROMT database (i.e. assigned to 7372400, Applications Software). This forced us to group several application markets together, leaving us with 29 software markets.

³⁷ This mean value is different from the number reported in Table 1, as we focus on years 2002 to 2009 across 29 software markets for this set of falsification tests.

systems and utilities, database software) consists of both an OSS and proprietary market, and assume the product entry events in both OSS or proprietary software market j in year t follows a Poisson process. The $X_{jt}'\beta$ in specification (1) should now be written as:

$$X_{jt}'\beta = \beta_1 Post-Commons_t * OSS-market_j + \gamma_1 PatentThicket_{jt-1} + \gamma_2 SalesGrowth_{jt} + \gamma_3 MarketPatents_{jt-1} + TimeTrend_{jt} \quad (4)$$

where *Post-Commons* is a dummy variable that is equal to one in 2005 and thereafter, and *OSS-market* is a dummy that is equal to one for OSS market. This regression is estimated using a conditional fixed effects Poisson model over a pooled data set with OSS and proprietary markets for each of the 29 different categories of software, i.e. with 58 separate markets ($j=58$ in specification (4)). *TimeTrend* is market-specific time trend, measured by a linear time trend times market dummies. We use market-specific time trend instead of baseline controls to address time-varying heterogeneity between OSS and proprietary markets. The results, reported in Table C-9 in the Appendix, suggest a greater increase in new product entry in OSS markets when compared with proprietary markets after 2005, though these effects do not increase significantly with cumulateness and concentration.

Table 7: Falsification test on the effect of The Commons on proprietary software entry, conditional fixed-effect Poisson regression

| Dependent variable: Proprietary software product entry | (1) | (2) | (3) | (4) |
|--|-----------------|----------------|-----------------|----------------|
| The Commons | .062 (.054) | .051 (.074) | .131 (.096) | .122 (.101) |
| The Commons * Cumulateness | | .017 (.094) | | 0.051 (.090) |
| The Commons * Concentration | | | -.182 (.244) | -.236 (.244) |
| Cumulateness | -3.566* (2.034) | -3.285 (2.890) | -3.941* (2.188) | -3.241 (2.900) |
| Concentration | -3.505 (6.283) | -3.671 (6.538) | -2.195 (6.593) | -2.292 (6.574) |
| Observations | 232 | 232 | 232 | 232 |
| Log pseudolikelihood | -443.038 | -442.993 | -442.390 | 442.066 |
| <i>Marginal Effects</i> | | | | |
| The Commons (average) | .062 (.059) | .062 (.054) | .090 (.059) | .100* (.056) |
| The Commons (cumulateness=10%) | | .054 (.064) | | .078 (.066) |
| The Commons (cumulateness=90%) | | .072 (.082) | | .131* (.081) |
| The Commons (concentration =10%) | | | .105 (.069) | .119* (.067) |
| The Commons (concentration =90%) | | | .075 (.051) | .080* (.051) |

Notes: 1) All regressions include year dummies and market fixed effects. 2) All regressions include other baseline controls such as sales growth, total patents, and patent quality. The complete results including estimates for the controls are reported in Appendix Table C-8. 3) Robust standard errors, clustered by market, are in parentheses. 4) Robust standard errors, clustered by market, are in parentheses. 5) * significant at 10%, ** significant at 5%, *** significant at 1%.

We next discuss the results of a series of instrumental variable (IV) estimates. Following the quasi-differencing method developed by Wooldridge (1997) and later employed by Windmeijer (2000) and Kim and Marschke (2005), we employ a count data model with IVs estimated using Generalized Method of Moments (GMM). Since 95% of the patents in The Commons were contributed by IBM in January 2005, our choice of IV mostly focuses on factors that influence IBM's decision on what to contribute and the timing of the contribution behavior.

Our first IV is the cumulated number of IBM's patents that were opposed at the European Patent Office (EPO) up to year t and related to software segment j (denoted as *IBM's opposed patents*). We

argue that because IBM’s strategic intent behind its contribution to The Commons is to support OSS innovation, its propensity to contribute in each market will be increasing in the extent to which it holds blocking patents across those markets. EPO oppositions will be a good proxy for the extent to which IBM holds blocking patents because of the institutional mechanisms through which oppositions are conducted. Different from US legal procedures, the EPO allows patents to be opposed up to nine months after the grant date and at a much lower cost than that of patents opposed through formal legal procedures. Third parties tend to oppose EPO granted patents that are more valuable, and oppositions are more likely to occur in areas characterized by cumulative innovation and uncertainty (Harhoff and Reitzig 2004).

The identification assumption for this IV is that IBM patents issued and opposed at the EPO are uncorrelated with OSS entry in the US market. This will be true when two conditions hold. First, among such EPO opposed patents, those with a related grant within the US are appropriately controlled for in the regression. Our variable *Total Patents* does include all relevant market patents, including patents granted to IBM in the US that may have been extended from or to the EPO. Second, industry conditions such as demand and technological opportunities driving EPO oppositions that are correlated across Europe and the US are similarly controlled for. While we have several controls for industry conditions, including *OSS Demand* and *Total Patents*, the controls are less than perfect and could still give rise to omitted variable biases in our estimates.

Our second instrument captures how changes in the perceived legal risks to developing OSS influenced the timing of IBM’s contribution. We interact IBM opposed patents with a time dummy that is equal to one in 2003 and thereafter (denoted as *afteryear2003*) to construct our second IV (denoted as *IBM’s opposed patents X afteryear2003*). The motivation to use *afteryear2003* dummy is driven by the SCO lawsuit against IBM on March 7, 2003. This was the first major IPR enforcement lawsuit targeting firms developing OSS-related product and attracted significant publicity. As such, it is expected to influence the timing of IBM’s patent contribution decision.

Table 8 shows the results of an auxiliary first stage OLS regression of the number of claims-weighted patents in The Commons on each of our instruments separately and then together.

Table 8: Results of OLS regression of Claims-weighted patents in The Commons on instruments

| Dependent variable: The Commons | IV: IBM’s opposed patents | IV: IBM’s opposed patents X afteryear2003 | The full set of IVs |
|---------------------------------------|---------------------------|---|---------------------|
| | (1) | (2) | (3) |
| IBM’s opposed patents | .226*** (.069) | | .115 (076) |
| IBM’s opposed patents X afteryear2003 | | .066*** (.016) | .051*** (.018) |
| First Stage F-statistic, p-value | 10.73, .001 | 17.53, .001 | 9.71, .001 |

Notes: 1) All regressions include other baseline controls such as sales growth, total patents, and patent quality. The complete results including estimates for the controls are reported in Appendix Table C-10. 2) The full set of IVs include IBM’s opposed patents and IBM’s opposed patents X afteryear2003. 3) The first stage OLS IV Regressions are used as auxiliary regressions to test for weak IVs, as there is no such test in using the GMM estimator. 4) Robust standard errors, clustered by market, are in parentheses. 5) * significant at 10%, ** significant at 5%, *** significant at 1%.

The direction of the coefficient estimates is as expected; the number of claims-weighted Commons patents is increasing in the number of IBM's patents opposed at the EPO, and this is particularly so after 2003. The joint F-test of instruments in the first stage shows the F-statistics ranging between 9.71 and 17.53, suggesting the IVs have some power in explaining the endogenous variable.

Table 9 below reports the GMM estimation results where we instrument for *The Commons* using all two IVs. As shown in column (1), the marginal effect of The Commons is significantly positive, consistent with our baseline estimates³⁸. The result from the Hansen J-statistic fails to reject the null that the instruments used are uncorrelated with the error term.

Table 9: GMM estimates of count data regression using the full set of IVs

| Dependent variable: OSS entry | Specification testing direct impact of The Commons only | Add interaction with cumulativeness of innovation | Add interaction with concentration of patent ownership | Add interaction with cumulativeness of innovation and with concentration of patent ownership |
|----------------------------------|---|---|--|--|
| | (1) | (2) | (3) | (4) |
| The Commons | .269* (.150) | .150 (.254) | -.157 (.261) | -.495** (.199) |
| The Commons * Cumulativeness | | .217 (.228) | | .497** (.199) |
| The Commons * Concentration | | | 1.315* (.808) | 1.550** (.800) |
| Cumulativeness | -4.644** (1.702) | -2.367 (2.639) | -3.029* (1.681) | .855 (1.305) |
| Concentration | .467 (4.366) | .336 (4.563) | -2.330 (5.178) | -2.686 (5.169) |
| Over-identification, p-value | .273 | .465 | .313 | .463 |
| <i>Marginal Effects</i> | | | | |
| The Commons (average) | .269* (.150) | .292** (.139) | .138 (.168) | .189** (.092) |
| The Commons (cumulativeness=10%) | | .193 (.211) | | -.038 (.088) |
| The Commons (cumulativeness=90%) | | .490*** (.183) | | .643*** (.231) |
| The Commons (concentration =10%) | | | .014 (.193) | .042 (.099) |
| The Commons (concentration =90%) | | | .282* (.182) | .358*** (.129) |

Notes: 1) All regressions include other baseline controls such as sales growth, total patents, and patent quality. The complete results including estimates for the controls are reported in Appendix Table C-11. 2) The full set of IVs include: IBM's opposed patents and IBM's opposed patents X afteryear2003. 2) Robust standard errors, clustered by market, are in parentheses. 3) * significant at 10%, ** significant at 5%, *** significant at 1%. 4) We use Wooldridge's quasi-differencing transformation to remove market fixed effects.

The remaining columns in Table 9 show the results of instrumental variables regressions in which the endogenous variables are *The Commons* interacted with *Cumulativeness* and *Concentration* and the instruments are the two instruments described above and their interactions with *Cumulativeness* and *Concentration*. Largely consistent with the baseline results, the estimated coefficients and implied marginal effects suggest a greater effect of The Commons when cumulativeness of innovation or patent ownership concentration is high; columns 3 and 4 show that the marginal effect of The Commons is significantly different when evaluated at high versus low level of cumulativeness and when evaluated at

³⁸ In this set of GMM estimations, we only use baseline controls (i.e. the controls used in column (1) in Table 4, as adding more controls tends to lead to non-convergent results.

high versus low level of concentration. For robustness, we also run GMM estimation using each of the two IVs individually. The results are in Table C-12 in the Appendix and are qualitatively similar to the results based on the full set of instruments.³⁹

7. Conclusions

In this study we provide the first large sample evidence of how the provision of a patent commons shapes OSS entry by start-ups. In particular, we examine the introduction of The Commons, a patent commons initiated by the Open Source Development Labs and IBM in 2005. We show that increases in the size of The Commons related to a software market are associated with increases in OSS entry by start-up software firms in that market. Furthermore, the impact of The Commons is magnified when two features of patent thickets are present in the market: cumulateness of innovation and concentration of patent ownership. We observe a particularly strong relationship between the size of The Commons and OSS start-ups' entry in markets with high cumulateness of innovation; the marginal effect of The Commons is also greater when concentration is high.

Our study has implications for our understanding of the impact of formal IPR on OSS, an area in which our knowledge is still quite limited (Lerner and Tirole 2005, von Hippel and von Krogh 2003). Despite the rapid growth in the commercialization of OSS,⁴⁰ OSS contributors and users are easily targeted by proprietary holders of IPR because of the cumulative nature of the development process and the difficulty in identifying the provenance of source code ownership. The existence of fear, uncertainty, and doubt (FUD) arising from the threat of litigation will be important for the commercialization of OSS. This research advances our understanding of how the existence and exercise of formal IPR influences the rate and direction of OSS innovation (Graham and Mowery 2005, Lerner and Tirole 2005, Maurer and Scotchmer 2006). This study also sheds some light on recent work that studies firm decisions to appropriate value from OSS (Bonaccorsi et al. 2006, Fosfuri et al. 2008).

Related, our research has important implications for technology management. Contributors to patent commons forego potential opportunities to license their IPR in hopes of increasing innovative activity that will spur demand for complementary products and services from which the contributor can appropriate value. If contributions do not lead to entry of new OSS products, then the benefits to contribution will be isolated to the positive publicity in the OSS community. Our results show that such contributions will have an effect on commercial activity in OSS, particularly in markets where patent thickets are most likely to raise ex ante entry costs.

³⁹ We have also implemented a Hausman test for the necessity of using GMM estimation with IVs. The chi-square obtained from comparing the estimates of our baseline Poisson model and GMM with IVs is 6.331, and insignificant, suggesting the estimates from the GMM with IVs do not reject the Poisson estimates.

⁴⁰ For example, a recent survey has indicated that worldwide revenue from commercializing OSS reached \$1.8 billion in 2006 and is expected to reach \$5.8 billion in 2011 (Broersma 2007).

Our results differ significantly from that in some recent research on traditional patent pools. Such pools allow firms to combine their patents as if they were a single firm, reducing transaction costs and licensing fees for a set of patented technologies. Different from the patent commons that we study, patent pools require some combination of donations to the pool and/or license fees for firms to use protected technologies. Recent research on patents pools has demonstrated that they can lead to a decline in innovation (Flamm 2013, Lampe and Moser 2010), and a change in the direction of innovation toward substitute technologies (Lampe and Moser 2013). One hypothesized reason for their result is differential licensing fees for access to patents for pool members and non-members. In contrast, the royalty-free patents pools that we study lead increases in the introduction of new commercial technologies in segments where commons patents are particularly concentrated.

Our findings also inform recent work on the impacts of patent thickets on start-up entry (e.g., Cockburn and MacGarvie 2011). Unlike large firms, start-ups usually lack the R&D capabilities and financial resources required to expand their own patent portfolios, so it is difficult for them to navigate patent thickets using other approaches such as cross-licensing agreements.⁴¹ However, start-ups may be offering complementary technologies within different platforms and may be unaware of the extent to which they are infringing patents or to which patents in the commons offer a means to avoid patent infringement. Our results are consistent with recent findings that the commercialization strategies of start-up firms are sensitive to the distribution of IPRs in the market (e.g., Gans et al. 2002, Huang et al. 2013).

We also contribute to the literature by examining the potential anti-commons problems from strategic patenting and the impact of patent thickets on entry into the software industry (e.g., Cockburn and MacGarvie 2011, Ziedonis 2004). While the patent thickets problem can be examined from different perspectives, we highlight the roles of cumulateness of innovation and patent ownership concentration as two different and important dimensions of patent thickets. We propose mechanisms under which these characteristics may interact with patent commons to determine start-up entry costs. Our research also provides empirical evidence that patent commons may be used as another mechanism, in addition to traditional patent pools or SSOs (Shapiro 2001, Rysman and Simcoe 2008), to effectively mitigate the anti-commons problem.

Our study is subject to limitations. One particular concern is how our results may be influenced by time-varying market unobservables that can affect both OSS entry and donations to The Commons. While we demonstrate that our results are robust to a range of analyses, we acknowledge that our empirical results may reflect different mechanisms through which The Commons may influence entry.

⁴¹ As noted by Matt Asay, the chief operating officer at Canonical (the company behind the Ubuntu Linux operating system), “this [type of patent collective] may be the only refuge for start-ups and others, like Red Hat, that don’t have an aggressive patent-acquisition policy.” (Matt Asay 2010)

First, the technologies contained in The Commons may directly reduce invention and licensing costs for start-ups. However, another mechanism is that through deploying a specific set of IPR, the contribution by IBM provides a signal for its broader efforts to protect OSS developers from litigation risks. We do not explicitly disentangle these different mechanisms, but we believe this is an important question for future study.

Our study also points to many other opportunities for future research. In particular, we analyze the impact of patent commons on the behavior of those firms whose entry decisions are most likely to be affected by the change in licensing and negotiation costs: start-up firms considering entry as an OSS competitor.⁴² The introduction of patent commons may have secondary implications for two groups of firms that we do not study: large firms and those who sell software under a traditional proprietary license. Understanding the implications of patent commons on these other groups will have important implications for the rate and direction of inventive activity in software. We hope our findings stimulate additional research in this important area.

References

- Alexy, O., M. G Reitzig. 2013. Private–collective Innovation, Competition, and Firms’ Counterintuitive Appropriation Strategies. *Research Policy* 42 (4): 895-913.
- Allison J. R., M. A. Lemley, K. A. Moore, R. D. Trunkey. 2004. Valuable Patents. *Georgetown Law Journal* 92: 435-479.
- Arora, A., A. Fosfuri, A. Gambardella. 2001. *Markets for Technology: Economics of Innovation and Corporate Strategy*. Cambridge: MIT Press.
- Asay, Matt. 2010. What Apple's and Microsoft's patent threats mean for start-ups. CNET. Available at http://news.cnet.com/8301-13505_3-10466670-16.html
- Auza, Jun. 2011. Infamous Microsoft FUD Campaigns Against Linux. Available at www.junauza.com, posted on July 18, 2011.
- Bessen, J. E., R. M. Hunt. 2007. An Empirical Look at Software Patents. *Journal of Economics & Management Strategy* 16(1): 157-89.
- Bessen, J. 2008. The value of U.S. patents by owner and patent characteristics. *Research Policy* 37(5): 932-945.
- Bessen, J., M. J. Meurer. 2008. *Patent Failure: How Judges, Bureaucrats, and Lawyers Put Innovators at Risk*. Princeton: Princeton University Press.
- Biddle, B., A. White, S. Woods. 2012. How Many Standards in a Laptop? (And Other Empirical Questions) Working Paper, Sandra Day O’Connor College of Law, Arizona State University.
- Bonaccorsi, A., S. Giannangeli, C. Rossi. 2006. Entry Strategies Under Competing Standards: Hybrid Business Models in the Open Source Software Industry. *Management Science* 52(7): 1085-1098.
- Broersma, M. 2007. IDC: Open-Source Market to Be Worth \$5.8B by 2011. CIO. Available at http://www.cio.com/article/116201/IDC_Open_Source_Market_to_Be_Worth_5.8B_by_2011
- Caballero, R.J., A.B. Jaffe. 1993. How High are the Giants' Shoulders: An Empirical Assessment of Knowledge Spillovers and Creative Destruction in a Model of Economic Growth. In O. Blanchard

⁴² A natural and interesting extension to this study is to look at how the size of The Commons influences the survival rate of start-up firms producing OSS products. However, of the firms introducing new OSS products in our sample, only one firm exited before the end of the sampling period. As a result, there is insufficient variance in our data to measure the impact of The Commons on the survival of start-up firms who produce OSS.

- and S. Fischer, eds., National Bureau of Economic Research Macroeconomics Annual, Vol. 8, pp. 15-86, MIT Press, Cambridge, MA.
- Campbell-Kelly, M., D.D. Garcia-Swartz. 2009. Pragmatism, not ideology-Historical perspectives on IBM's adoption of open-source software. *Information Economics and Policy*. 21(3): 229–244.
- Capek, P. G., S. P. Frank, S. Gerdt, and D. Shields. 2005. A history of IBM's open-source involvement and strategy. *IBM Systems Journal* 44(2): 249-257.
- Clarkson, G. 2005. Patent Informatics for Patent Thicket Detection: A Network Analytic Approach for Measuring the Density of Patent Space. Working Paper. Available at: <http://w4.stern.nyu.edu/emplibrary/ACFltbnmV.pdf>
- Coase, R. 1960. The Problem of Social Cost. *Journal of Law and Economics* 3(1): 1-44.
- Cockburn, I. M., M. MacGarvie. 2006. Entry, Exit and Patenting in the Software Industry. NBER Working paper 12563, National Bureau of Economic Research, Cambridge, MA.
- Cockburn, I. M., M. MacGarvie. 2009. Patents, Thickets and the Financing of Early-Stage Firms: Evidence from the Software Industry. *Journal of Economics & Management Strategy* 18(3): 729-773.
- Cockburn, I. M., M. MacGarvie. 2011. Entry and Patenting in the Software Industry. *Management Science* 57(5): 915-933.
- Cohen, J. E., M. A. Lemley. 2001. Patent scope and innovation in the software industry. *California Law Review*, 89(1): 1-57.
- Flamm, K. 2013. A Tale of Two Standards: Patent Pools and Innovation in the Optical Disk Drive Industry. *NBER Working Papers No 18931* National Bureau of Economic Research, Cambridge, MA
- Fosfuri, A., M. S. Giarratana, A. Luzzi. 2008. The Penguin Has Entered the Building: The Commercialization of Open Source Software Products. *Organization Science* 19(2): 292-305.
- Galasso, A., M. Schankerman. 2010. Patent Thickets, Courts, and the Market for Innovation. *RAND Journal of Economics* 41(3): 472-503.
- Gambardella, A., B. H. Hall. 2006. Proprietary versus public domain licensing of software and research products. *Research Policy* 35(6): 875-892.
- Gans, J. D. Hsu, and S. Stern. 2002. When does start-up innovation spur the gale of creative destruction? *RAND Journal of Economics* 33(4): 571-586.
- Graham, S. J. H., D. Mowery. 2005. The Use of USPTO "Continuation" Applications in the Patenting of Software: Implications for Free and Open Source. *Law & Policy* 27(1): 128-151.
- Green, J. R., S. Scotchmer. 1995. On the division of profit in sequential innovation. *The RAND Journal of Economics* 26(1): 20-33.
- Griliches, Z. 1990. Patent statistics as economic indicators: a survey. *Journal of Economic Literature* 27: 1661–1707.
- Hall, B.H., C. Helmers. 2013. Innovation and diffusion of clean/green technology: Can patent commons help? *Journal of Environmental Economics and Management* 66(1):33-51.
- Hall, B.H., C. Helmers, G. von Graevenitz, C. Rosazza-Bondibene. 2012. A Study of Patent Thickets: Final report prepared for the UK Intellectual Property Office.
- Hall B.H., A. Jaffe, M. Trajtenberg. 2005. Market value and patent citations. *RAND Journal of Economics* 36 (1): 16-38.
- Hall, B. H., R. H. Ziedonis. 2001. The patent paradox revisited: An empirical study of patenting in the U.S. semiconductor industry, 1979-1995. *RAND Journal of Economics* 32(1): 101-128.
- Harhoff, D., M. Reitzig. 2004. Determinants of opposition against EPO patent grants—the case of biotechnology and pharmaceuticals. *International Journal of Industrial Organization* 22(4): 443-480.
- Heller, M. A., R. S. Eisenberg. 1998. Can Patents Deter Innovation? The Anticommons in Biomedical Research. *Science* 280 (5364): 698-701.
- Huang, P., M. Ceccagnoli, C. Forman, D. J. Wu. 2013. Appropriability Mechanisms and the Platform Partnership Decision: Evidence from Enterprise Software. *Management Science* 59(1): 102–121.
- Jaffe, A.B, M.S. Fogarty, B.A. Banks. 1998. Evidence from Patents and Patent Citations on the Impact of NASA and Other Federal Labs on Commercial Innovation. *Journal of Industrial Economics* 46(2): 183-205.

- Jaffe, A. B., J. Lerner. 2004. *Innovation and Its Discontents: How Our Broken Patent System Is Endangering Innovation and Progress, and What to Do About It*. Princeton University Press, Princeton, NJ.
- Kim, J., G. Marschke. 2005. Labor Mobility of Scientists, Technological Diffusion, and the Firm's Patenting Decision. *RAND Journal of Economics* 36(2): 298-317.
- Lampe, R., P. Moser. 2010. Do Patent Commons Encourage Innovation? Evidence from the Nineteenth-Century Sewing Machine Industry. *The Journal of Economic History* 70(4): 898-920.
- Lampe, R., P. Moser. 2013. Patent Pools and Innovation in Substitutes - Evidence from the 19th Century Sewing Machine Industry. *RAND Journal of Economics*, forthcoming.
- Lanjouw, J. O., M. Schankerman. 2004. Patent Quality and Research Productivity: Measuring Innovation with Multiple Indicators. *The Economic Journal* 114(495): 441-465.
- Lemley, M. A., C. Shapiro. 2007. Patent Hold-Up and Royalty Stacking. Available at <http://escholarship.org/uc/item/8638s257>
- Lerner, J., J. Tirole. 2005. The Economics of Technology Sharing: Open Source and Beyond. *Journal of Economic Perspectives* 19 (Spring): 99-120.
- Lévêque, F., Ménière, Y. 2007. Copyright versus Patents: The Open Source Software Legal Battle. *Review of Economic Research on Copyright issues* 4(1): 27-46.
- Lichtman, D. 2006. Patent Holdouts in the Standard-Setting Process. U Chicago Law and Economics, Olin Working Paper No. 292.
- Mailath, G., A. Postlewaite. 1990. Asymmetric Information Bargaining Problems with Many Agents. *Review of Economic Studies* 57(3): 351-367.
- Maurer, S. M., S. Scotchmer. 2006. Open Source Software: The New Intellectual Property Paradigm. Hendershott, T eds. *Handbook of Economics and Information Systems* 285-319, Elsevier.
- Marson, Ingrid. 2004. Linus Torvalds speaks out against EU Patent Law. ZDNet. Available at <http://www.zdnet.com/linus-torvalds-speaks-out-against-eu-patent-law-3039174746/>. Accessed April 30, 2013.
- Noel, M.D., M.A. Schankerman. 2006. Strategic Patenting and Software Innovation: Theory and evidence from a panel of U.S. firms. CEPR Working Paper 5701, Center for Economic Policy Research, Washington, DC.
- Rysman, M., T. Simcoe. 2008. Patents and the Performance of Voluntary Standard Setting Organizations. *Management Science* 54(11): 1920-1934.
- Samuelson P. 2006. IBM's pragmatic embrace of open source. *Communications of the ACM*, 49(10): 21-25.
- Scannell, Ed. 2004. IBM says it won't assert patents against Linux kernel. InfoWorld Daily News, August 4, 2004.
- Schultz, J., J. M. Urban. 2012. Protecting Open Innovation: The Defensive Patent License as a New Approach to Patent Threats, Transaction Costs, and Tactical Disarmament. *Harvard Journal of Law and Technology* 26(1): 1-67.
- Scotchmer. 2004. *Innovation and Incentives*. MIT Press, Cambridge, MA.
- Serafino, D. 2007. Survey of Patent Commons Demonstrates Variety of Purposes and Management Structures. KEI Research Note 6.
- Shapiro, C. 2001. Navigating the Patent Thicket: Cross Licenses, Patent Commons, and Standard-Setting. A. Jaffe, J. Lerner, S. Stern, eds. *Innovation Policy and the Economy, Vol. 1*. MIT Press, Cambridge, MA.
- Silverman, B. S. 1999. Technological Resources and the Direction of Corporate Diversification: Toward an Integration of the Resource-Based View and Transaction Cost Economics. *Management Science* 45(8): 1109-1124.
- Simcoe, T. S., S. J. H. Graham, M. P. Feldman. 2009. Competing on Standards? Entrepreneurship, Intellectual Property, and Platform Technologies. *Journal of Economics & Management Strategy*. 18(3):775-816.

- Stallman, R. 2011. Europe's "unitary patent" could mean unlimited software patents. Available at <http://www.gnu.org/philosophy/europes-unitary-patent.html>. Access April 30, 2013.
- Thompson P., M. Fox-Kean. 2005. Patent Citations and the Geography of Knowledge Spillovers: A Reassessment. *The American Economic Review* 95(1): 450-460.
- von Graevenitz G., S. Wagner, D. Harhoff. 2011. How to measure patent thickets—A novel approach. *Economics Letters* 111(1): 6-9.
- von Hippel, E., G. von Krogh. 2003. Open Source Software and the 'Private-Collective' Innovation Model: Issues for Organization Science. *Organization Science* 14(2): 209-223.
- Windmeijer, F. 2000. Moment Conditions for Fixed Effects Count Data Models with Endogenous Regressors. *Economics Letters* 68(1): 21-24.
- Wooldridge, J.M. 1997. Multiplicative Panel Data Models without the Strict Exogeneity Assumption. *Econometric Theory* 13: 667-678.
- Ziedonis, R.H. 2004. Don't Fence Me In: Fragmented Markets for Technology and the Patent Acquisition Strategies of Firms *Management Science* 50(6): 804-820.

APPENDIX A: Patent Pledging Events

| Event Date | Pledging Firm(s) | Pledged Patent(s) | Potential licensees | Notes |
|------------|---|--|---|---|
| Jan 2005 | IBM | More than 500 specified patents (contributed to The Commons) | Anyone developing code under an OSI approved license | This pledging event has been included in our analysis |
| Jan 2005 | Sun Microsystems | 1670 (unspecified) patents related to Sun's Solaris | Developer working on any approved project under the Common Development and Distribution License (CDDL) | Two main criticisms of this pledge: 1. The CDDL doesn't permit mingling its code with code under GNU GPL, which governs Linux. This means developers can't use these patents on Linux – the freely granted patents can only enable idea-sharing among programmers for Solaris-related projects. 2. Sun's announcement was too broad and didn't specify these 1670 patents or respond to any developers' questions about what rights the developers have to these patents. |
| Sep 2005 | Computer Associates International Inc. | 14 patents (contributed to The Commons) | Anyone developing code under an OSI approved license | This pledging event has been included as a control in our analysis. |
| Nov 2005 | Nokia | Any of its patents | Developers working for the Linux Kernel only | Criticism: Because of Nokia's stance on Linux only, developers questioned why it did not apply to directly related projects such as GNOME and KDE and why it did not apply to application projects that are not necessarily directly related to Linux. |
| Nov 2005 | Open Invention Network, founded by IBM, Novell, Koninklijke Philips Electronics, Sony and Red Hat | Any of OIN's patents | Any company, institution or individual that agrees not to assert its patents against the Linux operating system or certain Linux-related applications | This pledging event has been included in our analysis as a control, as many of its patents have been pledged only recently and toward the end of our sample period. |
| Feb 2007 | Blackboard Inc. | Patent 6,988,138, 7,493,396, 7,558,853; pending patent applications: 12/470,739; 10/443,149; 10/643,075; 10/653,074; 11/142,965; 10/373,924; 10/918,016. | Anyone contributing to OSS projects, OSS initiatives, commercially developed open source add-on applications to proprietary products | For the commercially developed open source add-on applications to proprietary products, if the software's end license is open, then it is covered by the pledge; if it is partly open and partly proprietary, it is not covered. |

APPENDIX B: Identification of Software Markets and the Matching Patent Classes

Step 1: Identify Software Markets

To measure entry with new OSS products related to different software markets in each year, a crucial step is to divide the software market into different markets that are reasonably distinct from each other. One main source of software markets is the product code classification system embedded in the PROMT database. For a portion of news articles from PROMT, there are a few product codes assigned to each new article that indicate what product category/categories are associated with that article. All these product categories are organized as a hierarchical structure by PROMT and are defined both in terms of customer markets and technologies. Table B-1 shows some examples of PROMT codes.

However, there are two drawbacks to just relying on PROMT classifications. First, a significant percentage (about 60%) of OSS product introduction news articles from PROMT is missing the product code field. Thus, we must manually assign product codes for this set of articles. Second, the PROMT classes do not include keywords, making it difficult to manually match articles to PROMT classes. Thus, we further match PROMT product code classes with CorpTech product code classes⁴³ to take advantage of the keywords defined for each CorpTech product code. The resulting concordance table (denoted as the PROMT-CorpTech concordance hereafter) consists of about 80 PROMT codes matched to CorpTech's six-digit or seven-digit product codes. Each product code is associated with a set of technology phrases specific to that product code. This is used as a basis for us to identify (i) the PROMT articles with missing product codes and (ii) the related patents across a variety of software markets. Table B-2 shows some examples of the PROMT-CorpTech concordance.

Step 2: Identify Patent Classes across Software Markets

Using the NBER patent data project and USPTO database, we constructed our patent dataset, which consists of all patents granted from 1976 to 2009. Our sample period is from 1999 to 2009. To identify the related patents across a range of markets from the PROMT-CorpTech concordance, we first examined specialist firms that produce in only one CorpTech six- or seven-digit code⁴⁴. The sample of single specialists is from the CorpTech directory, over 1992 to 2004 and 2010.⁴⁵ We found 3500 patents held by about 700 specialists that have products associated with CorpTech codes from the PROMT-CorpTech concordance. The 3-digit USPTO classes to which the 3500 patents and their forward citations belong served as a starting point for us to map patent classes to each product code: for each product code, the top decile of these 3-digit US classes was used as candidates representing the core technologies for

⁴³ There are more than 290 software product codes (denoted as SOF) defined by CorpTech Directory. Each firm in this directory is associated with a set of self-reported product codes selected from these 290 SOF categories.

⁴⁴ Examples of CorpTech code are provided in Table B-2.

⁴⁵ Unfortunately, data from the CorpTech Directory from 2005 to 2009 was not available.

that code. While such endeavor to identify patent classes across software markets has been undertaken by prior research, in particular Cockburn and MacGarvie (2006, 2011), we constructed our own classification for several reasons. First, our sample period is more recent than theirs, so the mapping between patent technologies and product markets may have changed over time. Second, Cockburn and MacGarvie (2006, 2011) examined 28 specific CorpTech product codes that have incomplete overlap with the open source product markets that we study. Finally, we took the intersection of the patent classes from the patents in The Commons with the above mapping, which lead to 34 US patent classes and their corresponding product codes.

Step 3: Match Software Markets with Patent Class-subclass Combinations

Because most of the 3-digit US patent classes contain quite heterogeneous technologies, we then further generated a more detailed mapping between software product codes and US patent subclass levels by searching for technology phrases associated with each product code in patent class and subclass description. After identifying possible patent class-subclass, we then decided the concordance between product codes and patent class-subclass manually. This process generated the final mapping between software markets and patent class-subclass combinations. We further consolidated all product codes into 33 software markets based on whether they are supported by the same technologies (similar patent classes), as we are most interested in whether the supply of certain technologies by The Commons helps start-ups move into new technology area. The final concordance that we used in the empirical analysis consists of 33 software markets matched to 422 patent class-subclass combinations. Table B-3 shows some examples of this final concordance between software markets and US patent class-subclass combinations. To boost our confidence, we compared our concordance with the concordance by Cockburn and MacGarvie (2006, 2011). Among the 28 CorpTech codes examined by Cockburn and MacGarvie (2006, 2011), 18 of them closely relates to 12 software markets defined by us, though some of our software markets are broader. We compared the patent class-subclass combinations matched to the 18 product codes by Cockburn and MacGarvie (2006, 2011) with the patent class-subclass combinations matched to the 12 markets by us. As shown in Table B-4, overall, our concordance has good overlap with theirs. Figures B-1 and B-2 present a more concrete view on the above three steps.

Keywords used to identify OSS entry

We used the following set of keywords to search in PROMT news articles for introduction of software products that are licensed as open source. A software product is tagged as open source if it contains any of these keywords. We first implement automatic search and then manually check the results to ensure it is licensed as open source. Our choice of open source license terms is based on the distribution of open source licenses used by OSS projects at SourceForge.net, which is the largest

repository of OSS. Over 230,000 projects and over 3 million users and developers were registered before the end of year 2009 (SourceForge 2009).

| | |
|---|--|
| Keywords related to generic terms of OSS: | open source , open-sourced, OSS, FLOSS, source code, GPL-compatible, non-copyleft, copyleft, free software license, open source license, open-source license, public domain |
| Keywords related to open source licenses: | GPL, General Public License , GNU, Lesser General Public License, LGPL, BSD, FreeBSD, Apache License, Apache Software License, Artistic License, MIT License, Mozilla Public License |

Table B-1: Examples of PROMT Codes

| | |
|---------|-----------------------------------|
| 7372502 | Operating systems |
| 7372503 | Operating system enhancements |
| 7372504 | Graphical user interface software |
| 7372505 | Portable document software |
| 7372510 | Software development tools |
| 7372511 | CASE software |
| 7372512 | Programming utilities |
| 7372513 | Application development software |
| 7372514 | Debugging & testing software |
| 7372520 | Peripheral support software |
| 7372521 | Device driver software |
| 7372522 | Data acquisition software |
| 7372523 | Printer support software |
| 7372530 | Disk/file management software |

Table B-2: Examples of the PROMT-CorpTech Concordance

| CorpTech Code | PROMT Product Code |
|---------------|--|
| SOF-CS-F | 7372650 Fax software |
| SOF-DM-M | 7372421 DBMS |
| SOF-HL-M | 7372466 Medical practice software |
| SOF-ME-S | 7372544 Sound/audio software |
| SOF-OA-MB | 7372662 BBS software |
| SOF-OA-MC | 7372674 Videoconferencing software |
| SOF-OA-ME | 7372605 Electronic mail software |
| SOF-OA-MG | 7372630 Workgroup software |
| SOF-OA-P | 7372441 DTP software |
| SOF-TS-EC | 7372433 Civil engineering software |
| SOF-TS-EE | 7372434 Electrical engineering software |
| SOF-TS-ER | 7372423 Geographic information systems |
| SOF-UT-H | 7372521 Device driver software |
| SOF-UT-O | 7372561 Data center management software |
| SOF-UT-Q | 7372513 Application development software |
| SOF-UT-X | 7372691 Data encryption software |

Table B-3: Examples of the Concordance between Software Markets and US Patent Class-subclass Combinations

| Software Market | US class | Subclass Level 0 | Subclass Level 1 |
|----------------------------------|----------|--|--|
| Artificial Intelligence Software | 706 | Fuzzy Logic Hardware | Fuzzy Neural Network |
| Artificial Intelligence Software | 706 | Knowledge Processing System | Creation Or Modification |
| Artificial Intelligence Software | 706 | Knowledge Processing System | Knowledge Representation And Reasoning Technique |
| Artificial Intelligence Software | 706 | Neural Network | Learning Method |
| Artificial Intelligence Software | 706 | Neural Network | Learning Task |
| Artificial Intelligence Software | 706 | Neural Network | Neural Simulation Environment |
| Artificial Intelligence Software | 706 | Neural Network | Structure |
| Artificial Intelligence Software | 706 | Plural Processing Systems | |
| Data Encryption Software | 380 | Communication System Using Cryptography | Having Compression |
| Data Encryption Software | 380 | Communication System Using Cryptography | Time Market Interchange |
| Data Encryption Software | 380 | Facsimile Cryptography | Including Generation Of An Associated Coded Record |
| Data Encryption Software | 380 | Key Management | Having Particular Key Generator |
| Data Encryption Software | 380 | Key Management | Key Distribution |
| Data Encryption Software | 380 | Particular Algorithmic Function Encoding | NBS/DES Algorithm |
| Data Encryption Software | 380 | Particular Algorithmic Function Encoding | Public Key |
| Data Encryption Software | 380 | Video Cryptography | Copy Protection Or Prevention |
| Data Encryption Software | 726 | Access Control Or Authentication | Network |
| Data Encryption Software | 726 | Access Control Or Authentication | Stand-Alone |
| Data Encryption Software | 726 | Monitoring Or Scanning Of Software Or Data | |
| Data Encryption Software | 726 | Including Attack Prevention | Intrusion Detection |
| Data Encryption Software | 726 | Protection Of Hardware | Theft Prevention |

Note: 1) US patent class 706 is described as “Data processing: artificial intelligence”; US patent class 380 is described as “Cryptography”; US patent class 726 is described as “Information security”. 2) All subclasses within each US patent class are structured hierarchically. “Subclass level 0” means the subclass is on the highest level and “Subclass level 1” means the subclass is on the second highest level. Our mapping is based on subclass level 1.

Table B-4: Validation of market-patent concordance with Cockburn and MacGarvie (2006, 2011)

| # | Concordance by Cockburn and MacGarvie (2006, 2011) | | | | Our concordance | | |
|----|--|---|--------------|---------------------------|--|--------------|---|
| | Software code | Code description | Patent Class | Patent subclass | Software market | Patent class | Patent subclass |
| 1 | SOF-DM-MH | Software to present data to users in a tree-like structure | 707 | 1-10;100-104.1 | Database software | 707 | 1-10;100-102;104.1;200-206 |
| 2 | SOF-DM-MR | Software to present data to users in the form of tables of rows and columns | 707 | 1-10;100-104.1 | | | |
| 3 | SOF-DM-Q | Database query language software | 707 | 1-10;100-104.1 | | | |
| 4 | SOF-OA-ME | Electronic message systems software | 709 705 | 206; 008;009 | Electronic mail software | 709 | 206; |
| 5 | SOF-TS-ER | Geographic information systems software | 701 702 | 2XXX 005 | Geographic information systems | 701 | 117-223;300 |
| 6 | SOF-UT-H | Software to control the operation of peripheral devices | 710 | 1-74; | Peripheral device drivers | 710 719 | 8-19 321;324;328 |
| 7 | SOF-AC-B | Invoicing and billing software | 705 | 001-045 | Accounting, financial and business practice software | 705 | 16-45;51-59;64-80 |
| 8 | SOF-AC-T | Tax preparation and reporting software | 705 | 019;031 | | | |
| 9 | SOF-AI-A | Voice technology software | 704 | <=278 | Voice communications software | 704 | 200-256;256.1-256.8;257-278 |
| 10 | SOF-AI-L | Natural language software | 704 | 8;9 | Foreign language translation software | 704 | 1-10 |
| 11 | SOF-AI-N | Software that simulates the neural pathways of the brain | 706 | 15-44 | Artificial intelligence software | 706 | 2;10;12;14-62 |
| 12 | SOF-SV-AR | Artificial intelligence R&D | 706 | 15-62 | | | |
| 13 | SOF-DM-F | File management software | 707 | 1-10;200-206 | Disk/file management software | 711 707 | 100-173 1-10;200-206 |
| 14 | SOF-MA-Q | Software to control product quality | 700 | 108-115 | Manufacturing and business process software | 700 | 28-55;95-244;266-274 |
| 15 | SOF-WD-I | Inventory management software | 705 | 28;10 | | 705 | 7-11;28;29 |
| 16 | SOF-OA-GD | Three dimensional representation software | 345 700 | 418-427 98;118;119;120 | Graphics software | 345 | 418-428;440-443;467-475;502-506;519;520;536-689 |
| 17 | SOF-OA-P | Desktop publishing software | 715 | 500-542 | Personal productivity | 715 | 5XX |
| 18 | SOF-OA-W | Word processor/text editor software | 715 | 5XX | | | |
| | | | 707 | 200-206 | | | |

Figure B-1: Identification of Software Markets

| Software Market | PROMT Product Code | CorpTech Product Code |
|---|--|--|
| Database software | (NT5) 7372420 Database software (NT6) 7372421 DBMS (NT6) 7372422 DBMS utilities | SOF-DM (Database/file mgmt. software) Keywords: Database/file management software, DBMS, Relational DBMS, Information storage and retrieval systems software (ISRS) |
| Image analysis software | (NT5) 7372450 Image processing software (NT6) 7372459 Image processing software NEC | SOF-OA-GI (Image processing software) Keywords: Image processing software, Image analysis software, Image enhancement software |
| Manufacturing and business process software | (NT6) 7372414 Business information management software (NT6) 7372416 Manufacturing, distribution and retailing software | SOF-MA(Manufacturing software) Keywords: Manufacturing automation protocol software, Operations planning software, Manufacturing planning software, Process control manufacturing systems software, Software to control product quality, Production scheduling software |
| Software development tools | (NT5) 7372510 Software development tools (NT6) 7372511 CASE software (NT6) 7372512 Programming utilities (NT6) 7372513 Application development software (NT6) 7372514 Debugging & testing software | SOF-PD (Program development soft.) Keywords: Software development systems, Development environment sof, IDEs, Language compilers, Program translator, program translators, Cross assemblers SOF-UT-C (Debugging and testing soft.) Keywords: Debugging and testing software |

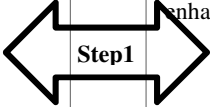
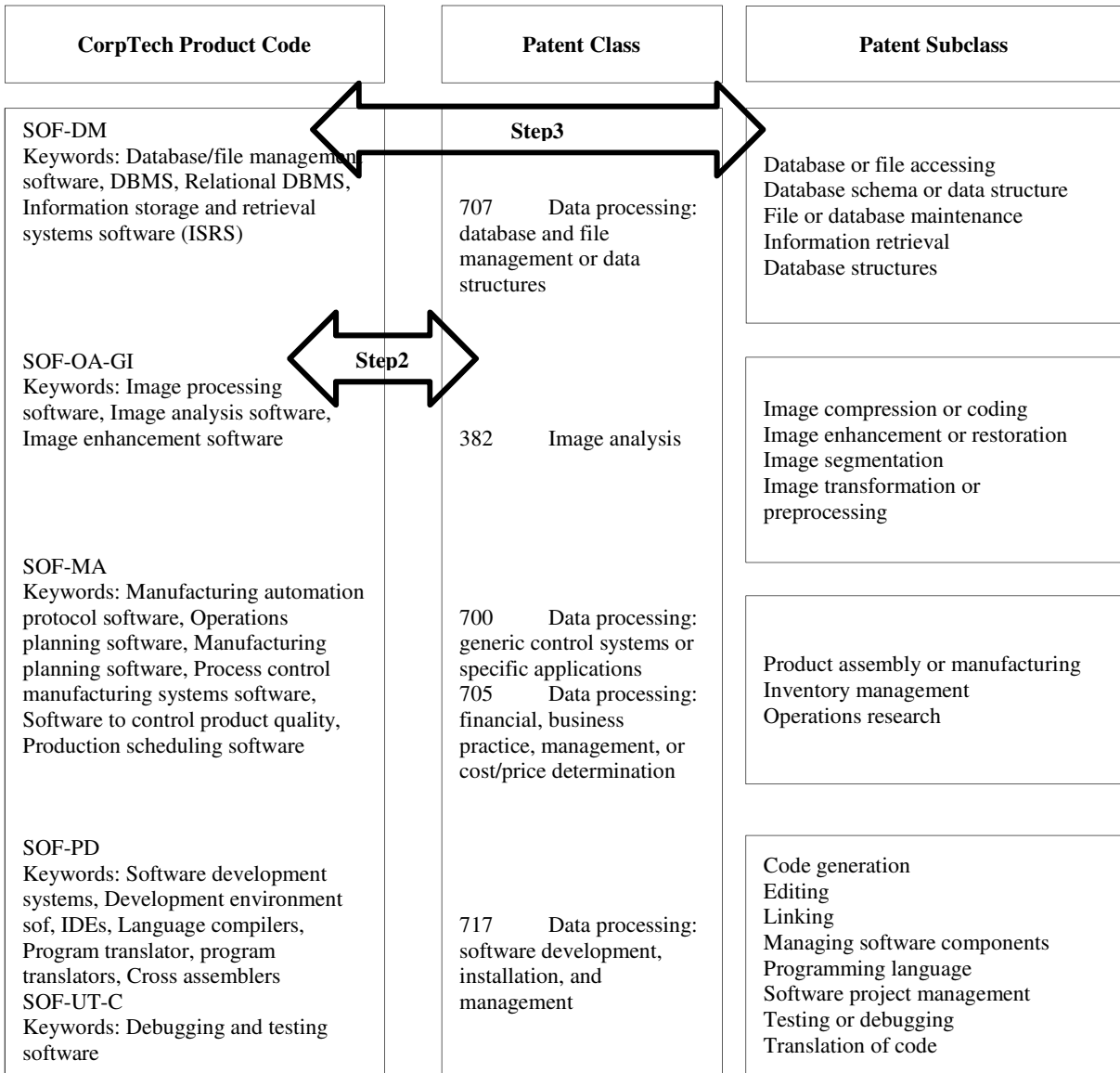


Figure B-2: Mapping Software Markets to Patent Subclasses



APPENDIX C: Supporting Empirical Results

Table C-1: Baseline results, linear fixed-effects regression

| Dependent variable: OSS entry | Specification testing direct impact of The Commons only | | Add interaction with cumulativeness of innovation | | Add interaction with concentration of patent ownership | | Add interaction with cumulativeness of innovation and with concentration of patent ownership | |
|---|---|------------------|---|------------------|--|------------------|--|-------------------|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| The Commons | .112* (.071) | .132* (.082) | .075 (.080) | .095 (.091) | -.091 (.082) | -.074 (.068) | -.121 (.091) | -.106 (.080) |
| The Commons * Cumulativeness | | | .112* (.061) | .111* (.066) | | | .106* (.062) | .107* (.064) |
| The Commons * Concentration | | | | | .724* (.405) | .730* (.384) | .708* (.396) | .719* (.375) |
| Cumulativeness | .381 (.672) | .583 (.780) | 1.235* (.727) | 1.400* (.787) | .673 (.621) | .810 (.729) | 1.475** (.726) | 1.591** (.791) |
| Concentration | .366 (3.722) | .243 (3.567) | .888 (3.570) | .742 (3.486) | 1.730 (3.593) | 1.596 (3.524) | 2.195 (3.445) | 2.053 (3.441) |
| Sales growth | -.479* (.316) | -.500 (.332) | -.515* (.317) | -.529* (.331) | -.552* (.354) | -.567* (.365) | -.584* (.352) | -.594* (.362) |
| Total patents | -.337 (.731) | -.193 (.829) | .214 (.777) | .310 (.839) | -.171 (.757) | -.098 (.798) | .346 (.752) | .384 (.775) |
| Patent quality | -.738 (1.179) | -.784 (1.201) | -.718 (1.210) | -.802 (1.237) | -.848 (1.166) | -.939 (1.180) | -.827 (1.198) | -.953 (1.217) |
| OIN patents | | -.026 (.037) | | -.020 (.036) | | -.033 (.035) | | -.027 (.034) |
| SSO patents | | -.037 (.074) | | -.037 (.071) | | -.023 (.058) | | -.024 (.056) |
| OSS demand | | .003 (.015) | | .009 (.016) | | .012 (.023) | | .018 (.022) |
| Year dummies | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Market fixed effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 363 | 363 | 363 | 363 | 363 | 363 | 363 | 363 |
| R-squared | 0.58 | 0.59 | 0.59 | 0.59 | 0.59 | 0.60 | 0.60 | 0.60 |
| <i>Marginal Effects</i> | | | | | | | | |
| The Commons (average) | .112* (.071) | .132* (.082) | .166** (.078) | .185** (.091) | .073* (.049) | .091* (.053) | .125*** (.047) | .143*** (.055) |
| The Commons (cumulativeness=10%) | | | .100 (.077) | .119 (.087) | | | .062 (.054) | .079 (.058) |
| The Commons (cumulativeness=90%) | | | .253** (.101) | .271** (.116) | | | .207*** (.073) | .226*** (.083) |
| Test of the difference between high and low cumulativeness, p-value | | | .067 | .091 | | | .087 | .094 |
| The Commons (concentration =10%) | | | | | .003 (.046) | .021 (.041) | .057 (.048) | .073* (.047) |
| The Commons (concentration =90%) | | | | | .151* (.079) | .170** (.085) | .201*** (.075) | .220*** (.083) |
| Test of the difference between high and low concentration, p-value | | | | | .074 | .057 | .073 | .055 |

Notes: 1) Robust standard errors, clustered by market, are in parentheses. 2) * significant at 10%, ** significant at 5%, *** significant at 1%. 3) R-squared includes fixed effects in R-squared computation.

Table C-2: Robustness test using raw number of patents, conditional fixed-effect Poisson regression

| Dependent variable: OSS entry | Specification testing direct impact of The Commons only | | Add interaction with cumulateness of innovation | | Add interaction with concentration of patent ownership | | Add interaction with cumulateness of innovation and with concentration of patent ownership | |
|--|---|-------------------|---|--------------------|--|--------------------|--|---------------------|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| The Commons | .040 (.139) | .086 (.152) | -.132 (.142) | -.080 (.150) | -.411 (.265) | -.564* (.297) | -.250 (.243) | -.282 (.316) |
| The Commons * Cumulateness | | | .686*** (.162) | .725*** (.170) | | | .645*** (.179) | .646*** (.208) |
| The Commons * Concentration | | | | | 1.255** (.600) | 1.824*** (.709) | .349 (.579) | .612 (.873) |
| Cumulateness | 1.613* (.939) | 1.646 (1.254) | 3.886*** (.832) | 4.357*** (.983) | 2.085** (.804) | 2.053** (1.004) | 3.890*** (.841) | 4.184*** (1.146) |
| Concentration | -5.203 (5.298) | -3.850 (6.315) | -10.581** (4.877) | -9.905* (6.375) | -10.325* (6.728) | -7.960 (6.866) | -11.664** (5.411) | -10.644* (6.325) |
| Sales growth | -.439 (.342) | -.446 (.368) | -.642* (.407) | -.690* (.458) | -.500 (.359) | -.551 (.401) | -.648* (.408) | -.700 (.464) |
| Total patents | .663 (2.030) | -.775 (2.287) | 2.113 (2.107) | 2.756 (2.270) | .790 (1.928) | .870 (2.073) | 2.076 (2.100) | 2.574 (2.290) |
| Patent quality | .109 (3.232) | -.095 (3.241) | .288 (3.165) | .333 (3.130) | .541 (3.195) | .596 (3.045) | .398 (3.120) | .524 (3.040) |
| OIN patents | | -.259 (.203) | | -.339** (.168) | | -.327* (.189) | | -.353** (.169) |
| SSO patents | | .038 (.197) | | -.009 (.143) | | .214 (.207) | | .054 (.200) |
| OSS demand | | .053 (.096) | | .020 (.093) | | .064 (.093) | | .028 (.093) |
| Year dummies | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Market fixed effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 286 | 286 | 286 | 286 | 286 | 286 | 286 | 286 |
| Log pseudolikelihood | -228.451 | -227.505 | -223.332 | -223.024 | -227.076 | -225.177 | -223.246 | -223.835 |
| <i>Marginal Effects</i> | | | | | | | | |
| The Commons (average) | .040 (.139) | .086 (.152) | .377*** (.144) | .457*** (.149) | -.133 (.163) | -.158 (.177) | .306* (.185) | .333* (.226) |
| The Commons (cumulateness=10%) | | | -.002 (.132) | .056 (.139) | | | -.051 (.153) | -.024 (.174) |
| The Commons (cumulateness=90%) | | | .879*** (.223) | .988*** (.229) | | | .778*** (.280) | .805** (.346) |
| Test of the difference between high and low cumulateness, p-value | | | .001 | .001 | | | .001 | .001 |
| The Commons (concentration =10%) | | | | | -.245 (.199) | -.322 (.220) | .275 (.223) | .278 (.291) |
| The Commons (concentration =90%) | | | | | -.005 (.136) | .029 (.147) | .342** (.152) | .395** (.166) |
| Test of the difference between high and low concentration, p-value | | | | | .036 | .009 | .547 | .483 |

Notes: 1) Robust standard errors, clustered by market, are in parentheses. 2) * significant at 10%, ** significant at 5%, *** significant at 1%. 3) The number of observations is lower than 363 because of the use of conditional fixed effects Poisson models, which drops markets without OSS entry over the entire sample period.

Table C-3: Robustness test using raw number of patents, linear fixed-effects regression

| Dependent variable: OSS entry | Specification testing direct impact of The Commons only | | Add interaction with cumulativeness of innovation | | Add interaction with concentration of patent ownership | | Add interaction with cumulativeness of innovation and with concentration of patent ownership | |
|---|---|------------------|---|--------------------|--|-------------------|--|-------------------|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| The Commons | .136 (.126) | .176 (.140) | .005 (.131) | .045 (.139) | -.322* (.166) | -.315* (.163) | -.373* (.178) | -.351* (.171) |
| The Commons * Cumulativeness | | | .388** (.177) | .401* (.200) | | | .319* (.176) | .319* (.188) |
| The Commons * Concentration | | | | | 1.633** (.786) | 1.758** (.811) | 1.427* (.760) | 1.515* (.771) |
| Cumulativeness | .406 (.656) | .698 (.834) | 1.957** (.784) | 2.426** (1.054) | .965* (.551) | 1.103 (.724) | 2.173** (.808) | 2.419** (.985) |
| Concentration | .897 (3.760) | .502 (3.681) | 2.255 (3.606) | 1.863 (3.583) | 1.738 (3.210) | 1.350 (3.320) | 2.751 (3.010) | 2.313 (3.172) |
| Sales growth | -.493* (.322) | -.544* (.347) | -.563* (.334) | -.622* (.361) | -.564* (.356) | -.612* (.377) | -.612* (.357) | -.664* (.379) |
| Total patents | -.132 (.976) | .212 (1.248) | 1.312 (1.296) | 1.832 (1.595) | .353 (1.021) | .473 (1.185) | 1.482 (1.251) | 1.723 (1.442) |
| Patent quality | -.434 (1.506) | -.347 (1.557) | .157 (1.630) | .301 (1.696) | -.217 (1.548) | -.296 (1.534) | .243 (1.617) | .212 (1.637) |
| OIN patents | | -.204* (.111) | | -.210** (.099) | | -.229** (.104) | | -.230** (.099) |
| SSO patents | | -.017 (.189) | | -.066 (.178) | | .054 (.139) | | .005 (.134) |
| OSS demand | | -.002 (.017) | | -.003 (.017) | | .011 (.021) | | .009 (.021) |
| Year dummies | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Market fixed effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 363 | 363 | 363 | 363 | 363 | 363 | 363 | 363 |
| R-squared | 0.52 | 0.52 | 0.53 | 0.53 | 0.53 | 0.53 | 0.54 | 0.54 |
| <i>Marginal Effects</i> | | | | | | | | |
| The Commons (average) | .136 (.126) | .176 (.140) | .318** (.145) | .369** (.171) | .048 (.095) | .083 (.103) | .208** (.087) | .250** (.103) |
| The Commons (cumulativeness=10%) | | | .089 (.119) | .132 (.130) | | | .020 (.096) | .062 (.102) |
| The Commons (cumulativeness=90%) | | | .620** (.252) | .682** (.298) | | | .456** (.197) | .498** (.222) |
| Test of the difference between high and low cumulativeness, p-value | | | .028 | .045 | | | .070 | .090 |
| The Commons (concentration =10%) | | | | | -.110 (.096) | -.086 (.096) | .071 (.095) | .104 (.098) |
| The Commons (concentration =90%) | | | | | .224 (.150) | .273* (.164) | .362*** (.137) | .413*** (.159) |
| Test of the difference between high and low concentration, p-value | | | | | .037 | .030 | .060 | .049 |

Notes: 1) Robust standard errors, clustered by market, are in parentheses. 2) * significant at 10%, ** significant at 5%, *** significant at 1%. 3) R-squared includes fixed effects in R-squared computation.

Table C-4: Robustness test using sample with different ending year, conditional fixed-effect Poisson regression

| Dependent variable: OSS entry | Sample ends at year 2008 | | | | Sample ends at year 2007 | | | |
|---|--------------------------|----------------------|---------------------|----------------------|--------------------------|----------------------|---------------------|----------------------|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| The Commons | .123 (.091) | .025 (.133) | -.162 (.137) | -.126 (.173) | .152 (.111) | .094 (.155) | -.127 (.155) | -.013 (.214) |
| The Commons * Cumulativeness | | .351*** (.096) | | .318*** (.101) | | .348*** (.099) | | .321*** (.109) |
| The Commons * Concentration | | | .770** (.381) | .411 (.405) | | | .716** (.361) | .258 (.409) |
| Cumulativeness | 1.004 (1.244) | 3.523*** (1.100) | 1.008 (1.113) | 3.296*** (1.141) | 1.030 (1.194) | 3.436*** (1.182) | 1.028 (1.100) | 3.277*** (1.193) |
| Concentration | -11.016 (8.088) | -16.289** (8.099) | -14.951* (8.432) | -17.811** (7.876) | -9.434 (7.568) | -15.160** (7.726) | -13.143* (8.012) | -15.989** (7.373) |
| Sales growth | -.418 (.398) | -.590 (.451) | -.506 (.418) | -.620 (.461) | -.533 (.354) | -.680 (.407) | -.616* (.359) | -.698* (.411) |
| Total patents | -.584 (1.559) | 1.111 (1.470) | -1.070 (1.585) | .716 (1.511) | -.204 (1.531) | 1.336 (1.502) | -.560 (1.526) | 1.127 (1.527) |
| Patent quality | -.851 (2.524) | -.645 (2.437) | -.820 (2.526) | -.631 (2.433) | .136 (2.785) | .316 (2.437) | .392 (2.785) | .397 (2.666) |
| OIN patents | -.032 (.071) | -.038 (.065) | -.040 (.069) | -.041 (.066) | -.012 (.080) | -.025 (.074) | -.016 (.080) | -.025 (.075) |
| SSO patents | -.020 (.081) | -.027 (.068) | .024 (.086) | -.003 (.082) | .013 (.084) | .009 (.069) | .052 (.087) | .022 (.082) |
| OSS demand | .082 (.093) | .073 (.080) | .084 (.098) | .074 (.084) | .090 (.091) | .079 (.082) | .094 (.095) | .081 (.084) |
| Year dummies | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Market fixed effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 260 | 260 | 260 | 260 | 225 | 225 | 225 | 225 |
| Log pseudolikelihood | -212.433 | -208.056 | -211.016 | -207.709 | -188.505 | -184.602 | -187.404 | -184.490 |
| <i>Marginal Effects</i> | | | | | | | | |
| The Commons (average) | .123 (.091) | .294** (.132) | .013 (.079) | .211* (.125) | .152 (.111) | .363** (.163) | .037 (.102) | .294* (.177) |
| The Commons (cumulativeness=10%) | | .095 (.128) | | .031 (.116) | | .166 (.154) | | .112 (.156) |
| The Commons (cumulativeness=90%) | | .556*** (.167) | | .449*** (.169) | | .634*** (.167) | | .544** (.231) |
| Test of the difference between high and low cumulativeness, p-value | | .001 | | .001 | | .001 | | .003 |
| The Commons (concentration =10%) | | | -.054 (.096) | .175 (.147) | | | -.027 (.118) | .271 (.202) |
| The Commons (concentration =90%) | | | .091 (.074) | .253** (.108) | | | .109 (.094) | .320** (.154) |
| Test of the difference between high and low concentration, p-value | | | .043 | .310 | | | .047 | .527 |

Notes: 1) Robust standard errors, clustered by market, are in parentheses. 2) * significant at 10%, ** significant at 5%, *** significant at 1%.

Table C-5: Robustness test adding “Top 4 incumbents market share”, conditional fixed-effect Poisson regression

| Dependent variable: OSS entry | (1) | (2) | (3) | (4) | (5) | (6) |
|--|--------------------|---------------------|-------------------|--------------------|---------------------|---------------------|
| The Commons | -.113 (.140) | -.088 (.170) | .199 (.128) | .122 (.159) | -.069 (.148) | -.048 (.174) |
| The Commons * Cumulativeness | | .305*** (.101) | | .302*** (.097) | | .294*** (.097) |
| The Commons * Concentration | .665* (.361) | .344 (.375) | | | .946** (.447) | .601 (.442) |
| The Commons * Top 4 incumbent market share | | | -.121 (.150) | -.147 (.140) | -.214 (.147) | -.188 (.143) |
| Cumulativeness | 1.078 (1.032) | 3.195*** (1.015) | 1.079 (1.218) | 3.153*** (.878) | 1.192*** (1.044) | 3.209*** (1.022) |
| Concentration | -10.542 (7.553) | -13.167* (6.951) | | | -9.879 (7.256) | -12.543* (6.726) |
| Sales growth | -.598 (.423) | -.716* (.408) | -.405 (.508) | -.579 (.497) | -.311 (.465) | -.441 (.467) |
| Total patents | -1.066 (1.578) | .624 (1.516) | -.133 (1.538) | 1.570 (1.341) | .973 (1.578) | .633 (1.466) |
| Patent quality | -1.731 (2.422) | -1.438 (2.342) | -1.546 (2.347) | -1.190 (2.245) | -1.965 (2.418) | -1.672 (2.310) |
| Top 4 incumbents market share | 1.294 (1.277) | 1.280 (1.129) | 1.483 (1.283) | 1.574 (1.112) | 1.649 (1.225) | 1.549 (1.139) |
| OIN patents | -.080 (.069) | -.076 (.066) | -.069 (.071) | -.072 (.066) | -.079 (.069) | -.075 (.065) |
| SSO patents | .027 (.083) | .004 (.078) | .003 (.064) | .009 (.050) | .033 (.084) | .010 (.078) |
| OSS demand | .062 (.099) | .060 (.086) | .063 (.091) | .055 (.078) | .063 (.095) | .058 (.082) |
| Year dummies | Yes | Yes | Yes | Yes | Yes | Yes |
| Market fixed effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 286 | 286 | 286 | 286 | 286 | 286 |
| Log pseudolikelihood | -226.032 | -222.856 | -227.163 | -223.643 | -224.945 | -222.031 |
| The Commons (average) | .034 (.086) | .215* (.123) | .151* (.092) | .287** (.130) | .054 (.085) | .227* (.121) |
| The Commons (cumulativeness=10%) | | .046 (.118) | | .120 (.125) | | .064 (.115) |
| The Commons (cumulativeness=90%) | | .438*** (.162) | | .508*** (.165) | | .442*** (.159) |
| Test of the difference between high and low cumulativeness, p-value | | .003 | | .002 | | .003 |
| The Commons (concentration =10%) | -.025 (.103) | .184 (.143) | | | -.030 (.107) | .173 (.141) |
| The Commons (concentration =90%) | .102 (.079) | .250** (.108) | | | .150* (.079) | .288** (.112) |
| Test of the difference between high and low concentration, p-value | .066 | .358 | | | .034 | .174 |
| The Commons (Top 4 incumbent market share =10%) | | | .181* (.112) | .324** (.151) | .107 (.094) | .274** (.132) |
| The Commons (Top 4 incumbent market share =90%) | | | .118 (.084) | .248** (.115) | -.003 (.092) | .177 (.120) |
| Test of the difference between high and low top 4 incumbents market share, p-value | | | .421 | .295 | .148 | .189 |

1) Robust standard errors, clustered by market, are in parentheses. 2) * significant at 10%, ** significant at 5%, *** significant at 1%.

Table C-6: Robustness test adding “Top 4 incumbents market share”, linear fixed-effects regression

| Dependent variable: OSS entry | (1) | (2) | (3) | (4) | (5) | (6) |
|---|------------------|-------------------|------------------|------------------|-------------------|-------------------|
| The Commons | -.075 (.066) | -.106 (.078) | .112 (.097) | .072 (.112) | -.055 (.079) | -.086 (.090) |
| The Commons * Cumulativeness | | .103* (.065) | | .106* (.065) | | .098 (.065) |
| The Commons * Concentration | .700* (.383) | .693* (.375) | | | .887** (.399) | .863** (.399) |
| The Commons * Top 4 incumbent market share | | | .012 (.178) | .022 (.176) | -.128 (.166) | -.116 (.167) |
| Cumulativeness | .828 (.733) | 1.581** (.778) | .625 (.774) | 1.402* (.772) | .921 (.711) | 1.625** (.779) |
| Concentration | 1.678 (3.650) | 2.108 (3.533) | | | 2.567 (3.448) | 2.893 (3.380) |
| Sales growth | -.619* (.373) | -.637* (.373) | -.589 (.473) | -.618 (.472) | -.482 (.466) | -.512 (.467) |
| Total patents | -.024 (.843) | .431 (.808) | -.115 (.689) | .317 (.682) | .025 (.822) | .451 (.800) |
| Patent quality | -.874 (1.183) | -.898 (1.221) | -.714 (1.188) | -.757 (1.224) | -1.098 (1.292) | -1.099 (1.326) |
| Top 4 incumbents market share | .981 (.922) | .834 (.879) | 1.366 (.846) | 1.188 (.789) | 1.108 (.839) | .957 (.794) |
| OIN patents | -.031 (.034) | -.026 (.033) | -.024 (.036) | -.019 (.035) | -.034 (.035) | -.029 (.034) |
| SSO patents | -.021 (.059) | -.021 (.057) | -.033 (.078) | -.034 (.076) | -.012 (.064) | -.014 (.062) |
| OSS demand | .011 (.023) | .017 (.023) | .002 (.017) | .009 (.018) | .009 (.022) | .014 (.022) |
| Year dummies | Yes | Yes | Yes | Yes | Yes | Yes |
| Market fixed effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 363 | 363 | 363 | 363 | 363 | 363 |
| R-square | .527 | .529 | .521 | .523 | .529 | .529 |
| The Commons (average) | .084* (.053) | .134** (.056) | .116* (.073) | .166** (.083) | .098* (.054) | .144** (.062) |
| The Commons (cumulativeness=10%) | | .074 (.059) | | .104 (.082) | | .087 (.061) |
| The Commons (cumulativeness=90%) | | .215** (.085) | | .249** (.108) | | .221** (.092) |
| Test of the difference between high and low cumulativeness, p-value | | .109 | | .103 | | .135 |
| The Commons (concentration =10%) | .017 (.041) | .068 (.048) | | | .012 (.047) | .061 (.054) |
| The Commons (concentration =90%) | .159* (.086) | .209** (.085) | | | .193** (.084) | .237 (.090) |
| Test of the difference between high and low concentration, p-value | .067 | .064 | | | .026 | .031 |
| The Commons (Top 4 incumbent market share =10%) | | | .114 (.081) | .161* (.092) | .126* (.068) | .170** (.077) |
| The Commons (Top 4 incumbent market share =90%) | | | .120 (.091) | .172* (.097) | .061 (.068) | .111* (.070) |
| Test of the difference between high and low top 4 incumbent market share, p-value | | | .947 | .903 | .440 | .486 |

1) Robust standard errors, clustered by market, are in parentheses. 2) * significant at 10%, ** significant at 5%, *** significant at 1%.

Table C-7: Complete results of falsification test by adding “False Commons”, conditional fixed-effect Poisson regression

| Dependent variable: OSS entry | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---------------------|----------------------|----------------------|----------------------|----------------------|-----------------------|
| The Commons | .373*** (.117) | .168 (.178) | .383*** (.140) | .093 (.164) | .426*** (.134) | .209 (.172) |
| False Commons | .463*** (.137) | .623** (.282) | | | .260 (.200) | .506 (.324) |
| IBM patents | | | .908*** (.249) | .867*** (.203) | .909*** (.257) | .855*** (.202) |
| IBM patents X 2003_2004 | | | .565*** (.160) | .504*** (.184) | .308 (.251) | .297 (.404) |
| The Commons X Cumulativeness | | .207 (.132) | | .076 (.121) | | .120 (.136) |
| False Commons X Cumulativeness | | .026 (.136) | | | | .077 (.208) |
| The Commons X Concentration | | .302 (.470) | | .646 (.405) | | .423 (.465) |
| False Commons X Concentration | | -.531 (.670) | | | | -.638 (.703) |
| Cumulativeness | -.272 (1.174) | 1.371 (1.402) | -.145 (1.113) | .497 (1.217) | -.413 (1.061) | .546 (1.420) |
| Concentration | -12.659* (6.569) | -17.654** (7.055) | -14.613** (7.245) | -17.900** (6.933) | -15.932** (6.949) | -20.170*** (7.029) |
| Sales growth | -.547 (.421) | -.704 (.462) | -.501 (.375) | -.596 (.395) | -.492 (.389) | -.607 (.417) |
| Total patents | -2.394 (1.516) | -1.393 (1.560) | -2.657* (1.485) | -2.539* (1.508) | -3.051** (1.518) | -2.655* (1.580) |
| Patent quality | -3.259 (2.334) | -2.953 (2.370) | -2.656 (2.192) | -2.527 (2.174) | -3.129 (2.310) | -3.103 (2.350) |
| OIN patents | -.055 (.069) | -.064 (.066) | -.050 (.064) | -.059 (.064) | -.052 (.066) | -.060 (.064) |
| SSO patents | -.021 (.071) | -.013 (.083) | -.001 (.071) | .038 (.080) | -.008 (.068) | .007 (.084) |
| OSS demand | .047 (.099) | .047 (.096) | .156* (.093) | .153* (.094) | .158* (.096) | .154* (.093) |
| Year dummies | Yes | Yes | Yes | Yes | Yes | Yes |
| Market fixed effects | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 286 | 286 | 286 | 286 | 286 | 286 |
| Log pseudolikelihood | -222.405 | -219.586 | -219.448 | -218.066 | -218.754 | -216.825 |
| <i>Marginal effects</i> | | | | | | |
| The Commons (average) | .373*** (.117) | .389** (.119) | .383*** (.140) | .292*** (.121) | .426*** (.134) | .392*** (.124) |
| The Commons (cumulativeness=10%) | | .275** (.127) | | .250** (.127) | | .326** (.128) |
| The Commons (cumulativeness=90%) | | .540*** (.168) | | .347** (.162) | | .480*** (.177) |
| Test of the difference between high and low cumulativeness, p-value | | .118 | | .531 | | .377 |
| The Commons (concentration =10%) | | .362** (.140) | | .234* (.138) | | .355** (.144) |
| The Commons (concentration =90%) | | .420*** (.111) | | .358*** (.113) | | .435*** (.116) |
| Test of the difference between high and low concentration, p-value | | .520 | | .110 | | .363 |

Notes: 1) Robust standard errors, clustered by market, are in parentheses. 2) * significant at 10%, ** significant at 5%, *** significant. 3) *False Commons* is measured by interacting the patent density in The Commons across markets with a time dummy that turns on for year 2003 and 2004. 4) *2003_2004* is a dummy that turns on for years 2003 and 2004. 5) *IBM patents* is measured by log of claims-weighted patent count of all IBM patents (excluding IBM patents contributed to The Commons) related to market *j* cumulated up to year *t*.

Table C-8: Complete results of falsification test on the effect of The Commons on proprietary software entry, conditional fixed-effect Poisson regression

| Dependent variable: Proprietary software product entry | (1) | (2) | (3) | (4) |
|---|--------------------|-------------------|--------------------|--------------------|
| The Commons | .062 (.054) | .051 (.074) | .131 (.096) | .122 (.101) |
| The Commons * Cumulativeness | | .017 (.094) | | 0.051 (.090) |
| The Commons * Concentration | | | -.182 (.244) | -.236 (.244) |
| Cumulativeness | -3.566* (2.034) | -3.285 (2.890) | -3.941* (2.188) | -3.241 (2.900) |
| Concentration | -3.505 (6.283) | -3.671 (6.538) | -2.195 (6.593) | -2.292 (6.574) |
| Sales growth | .193 (.247) | .197 (.259) | .194 (.241) | .206 (.252) |
| Total patents | 5.055* (2.796) | 5.099* (2.808) | 5.866* (3.076) | 6.234** (3.152) |
| Patent quality | 7.730* (5.160) | 7.632* (5.277) | 8.736* (5.446) | 8.754* (5.514) |
| OIN patents | 0.063 (.055) | .063* (.055) | .056 (.053) | .056 (.053) |
| SSO patents | -.073 (.065) | -.074 (.068) | -.080 (.066) | -.086 (.070) |
| OSS demand | -.204 (.191) | -.206 (.192) | -.239 (.216) | -.254 (.222) |
| Year dummies | Yes | Yes | Yes | Yes |
| Market fixed effects | Yes | Yes | Yes | Yes |
| Observations | 232 | 232 | 232 | 232 |
| Log pseudolikelihood | -443.038 | -442.993 | -442.390 | 442.066 |
| <i>Marginal Effects</i> | | | | |
| The Commons (average) | .062 (.059) | .062 (.054) | .090 (.059) | .100* (.056) |
| The Commons (cumulativeness=10%) | | .054 (.064) | | .078 (.066) |
| The Commons (cumulativeness=90%) | | .072 (.082) | | .131* (.081) |
| Test of the difference between high and low cumulativeness, p-value | | .852 | | .571 |
| The Commons (concentration =10%) | | | .105 (.069) | .119* (.067) |
| The Commons (concentration =90%) | | | .075 (.051) | .080* (.051) |
| Test of the difference between high and low concentration, p-value | | | .457 | .334 |

Notes: 1) We use the press releases of the 2,054 firms in the PROMT database to identify new proprietary software product entry. To identify products related to each market, we focus only on introduction events associated with PROMT product codes. For each start-up, we only use the firm's first product in a market to capture entry. This results in 2,384 proprietary product entry events from 2002 to 2009. We also adjust for a change in the assignment of product codes during our sample. Specifically, between 2007 and 2009 application-related software products were systematically assigned to a higher product code level in the PROMT database (i.e. assigned to 7372400, Applications Software). This forced us to group several application markets together, leaving us with 29 software markets. 2) Robust standard errors, clustered by market, are in parentheses. 3) * significant at 10%, ** significant at 5%, *** significant at 1%.

Table C-9: Falsification test on product entry using market-specific time trends, conditional fixed-effect Poisson regression

| Dependent variable: Product entry by start-ups | Specification testing direct impact of Post-Commons* OSS-market | | Add interaction with cumulativeness of innovation | | Add interaction with concentration of patent ownership | | Add interaction with cumulativeness of innovation and with concentration of patent ownership | |
|---|---|----------------------|---|----------------------|--|----------------------|--|----------------------|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Post-Commons* OSS-market | 2.021*** (.317) | 1.948*** (.330) | 1.910*** (.621) | 1.720*** (.660) | 2.076 (1.849) | 1.557 (1.727) | 1.781 (1.790) | 1.613 (1.745) |
| Post-Commons* OSS-market* Cumulativeness | | | .424 (1.209) | .560 (1.223) | | | .026 (1.174) | .515 (1.251) |
| Post-Commons* OSS-market* Concentration | | | | | -.192 (6.446) | 1.584 (6.198) | 1.206 (6.348) | .507 (6.401) |
| Cumulativeness | | 16.818 (10.552) | 5.634 (5.869) | 16.871* (10.562) | | 16.819 (10.548) | 5.415 (6.145) | 16.867 (10.565) |
| Concentration | | 28.008*** (7.925) | | 27.958*** (7.938) | 20.926* (12.827) | 27.938*** (7.927) | 20.067* (11.668) | 27.940*** (7.923) |
| Sales growth | | -.051 (.523) | | -.052 (.523) | | -.051 (.523) | | -.052 (.523) |
| Total patents | | 3.696 (6.976) | | 3.703 (6.976) | | 3.684 (6.980) | | 3.698 (6.983) |
| Patent quality | | -2.731 (5.400) | | -2.736 (5.403) | | -2.743 (5.405) | | -2.739 (5.407) |
| Market-specific time trend | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Market fixed effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 320 | 320 | 320 | 320 | 320 | 320 | 320 | 320 |
| Log pseudolikelihood | -722.515 | -680.232 | -717.518 | -680.170 | -713.354 | -680.212 | -708.919 | -680.169 |
| <i>Marginal Effects</i> | | | | | | | | |
| Post-Commons* OSS-market (average) | 2.021*** (.317) | 1.948*** (.330) | 2.173*** (.391) | 2.068*** (.371) | 2.032*** (.405) | 1.925*** (.383) | 2.077*** (.450) | 2.051*** (.502) |
| Post-Commons* OSS-market (cumulativeness=10%) | | | 1.988*** (.446) | 1.823*** (.476) | | | 2.066*** (.451) | 1.826*** (.461) |
| Post-Commons* OSS-market (cumulativeness=90%) | | | 2.424** (.990) | 2.400** (.964) | | | 2.093** (1.071) | 2.356** (1.121) |
| Test of the difference between high and low cumulativeness, p-value | | | .726 | .647 | | | .982 | .680 |
| Post-Commons* OSS-market (concentration =10%) | | | | | 2.047** (.897) | 1.795** (.823) | 1.978** (.948) | 2.009** (.961) |
| Post-Commons* OSS-market (concentration =90%) | | | | | 2.015*** (.313) | 2.063*** (.388) | 2.182*** (.389) | 2.095*** (.381) |
| Test of the difference between high and low concentration, p-value | | | | | .976 | .798 | .849 | .937 |

Notes: 1) Robust standard errors, clustered by market, are in parentheses. 2) * significant at 10%, ** significant at 5%, *** significant at 1%. 3) While we expand our original 29 software markets to 29 OSS markets and 29 proprietary software markets, 18 markets are automatically dropped because of no entry observed throughout the sampling years, which gives us 320 observations in total. 4) These results need to be interpreted with caution, as in this test we only used entry events (for both OSS and proprietary products) identified through PROMPT articles assigned with product codes (as mentioned earlier, on average 45% of product introduction events miss product codes).

Table C-10: Complete results of OLS regression of claims-weighted patents in The Commons on instruments

| Dependent variable: The Commons | IV: IBM's opposed patents | IV: IBM's opposed patents X afteryear2003 | The full set of IVs |
|---------------------------------------|---------------------------|---|---------------------|
| | (1) | (2) | (3) |
| IBM's opposed patents | .226*** (.069) | | .115 (076) |
| IBM's opposed patents X afteryear2003 | | .066*** (.016) | .051*** (.018) |
| Cumulativeness | 5.383*** (.523) | 5.388*** (.523) | 5.252*** (.526) |
| Concentration | -6.818** (3.180) | -6.392** (2.922) | -7.045** (2.985) |
| Sales growth | .138 (.285) | .242 (.280) | .186 (.282) |
| Total patents | 4.857*** (.749) | 4.507*** (.745) | 4.560*** (.749) |
| Patent quality | 3.775*** (1.184) | 2.908** (1.221) | 3.153** (1.242) |
| Year dummies | Yes | Yes | Yes |
| Market fixed effects | Yes | Yes | Yes |
| First Stage F-statistic, p-value | 10.73, .001 | 17.53, .001 | 9.71, .001 |

Notes: 1) The full set of IVs include IBM's opposed patents and IBM's opposed patents X afteryear2003. 2) The first stage OLS IV Regressions are used as auxiliary regressions to test for weak IVs, as there is no such test in using the GMM estimator. 3) Robust standard errors, clustered by market, are in parentheses. 4) * significant at 10%, ** significant at 5%, *** significant at 1%.

Table C-11: Complete results of GMM estimates of count data regression using the full set of IVs

| Dependent variable: OSS entry | Specification testing direct impact of The Commons only | Add interaction with cumulateness of innovation | Add interaction with concentration of patent ownership | Add interaction with cumulateness of innovation and with concentration of patent ownership |
|--|---|---|--|--|
| | (1) | (2) | (3) | (4) |
| The Commons | .269* (.150) | .150 (.254) | -.157 (.261) | -.495** (.199) |
| The Commons * Cumulateness | | .217 (.228) | | .497** (.199) |
| The Commons * Concentration | | | 1.315* (.808) | 1.550** (.800) |
| Cumulateness | -4.644** (1.702) | -2.367 (2.639) | -3.029* (1.681) | .855 (1.305) |
| Concentration | .467 (4.366) | .336 (4.563) | -2.330 (5.178) | -2.686 (5.169) |
| Sales growth | 1.140** (.478) | .736 (.509) | .500 (.556) | .011 (.723) |
| Total patents | -1.327* (.786) | -.194 (1.523) | -.728 (.785) | 1.505** (.746) |
| Patent quality | 3.430*** (1.173) | 2.963*** (1.024) | 2.049* (1.203) | 1.815 (1.137) |
| Year dummies | Yes | Yes | Yes | Yes |
| Observations | 363 | 363 | 363 | 363 |
| Over-identification, p-value | .273 | .465 | .313 | .463 |
| <i>Marginal Effects</i> | | | | |
| The Commons (average) | .269* (.150) | .292** (.139) | .138 (.168) | .189** (.092) |
| The Commons (cumulateness=10%) | | .193 (.211) | | -.038 (.088) |
| The Commons (cumulateness=90%) | | .490*** (.183) | | .643*** (.231) |
| Test of the difference between high and low cumulateness, p-value | | .341 | | .012 |
| The Commons (concentration =10%) | | | .014 (.193) | .042 (.099) |
| The Commons (concentration =90%) | | | .282* (.182) | .358*** (.129) |
| Test of the difference between high and low concentration, p-value | | | .103 | .050 |

Notes: 1) The full set of IVs include IBM's opposed patents and IBM's opposed patents X afteryear2003. 2) Robust standard errors, clustered by market, are in parentheses. 3) * significant at 10%, ** significant at 5%, *** significant at 1%. 4) We use Wooldridge's quasi-differencing transformation to remove market fixed effects.

Table C-12: GMM estimates of count data regression using each subset of IVs

| Dependent variable: OSS entry | IV: IBM's opposed patents | | | | IV: IBM's opposed patents X afteryear2003 | | | |
|----------------------------------|---------------------------|--------------------|-------------------|--------------------|---|---------------------|--------------------|-------------------|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| The Commons | .223 (.165) | -.020 (.278) | -.845 (.707) | -.483 (.324) | .235 (.154) | .088 (.247) | -.900 (.715) | -.512* (.274) |
| The Commons * Cumulativeness | | .374 (.266) | | .476** (.239) | | .278 (.272) | | .572** (.265) |
| The Commons * Concentration | | | 2.474* (1.301) | 1.881** (.911) | | | 2.379** (1.154) | 1.692** (.807) |
| Sales growth | 1.056** (.462) | .592 (.540) | .216 (.446) | -.065 (.555) | 1.188** (.555) | .942* (.574) | .427 (.618) | -.208 (.843) |
| Cumulativeness | -4.415*** (1.519) | -1.602 (2.895) | -2.942 (2.852) | -1.531 (2.853) | -5.196*** (1.750) | -2.871 (2.731) | -2.820 (2.501) | -.162 (2.762) |
| Concentration | .647 (4.233) | 1.172 (4.665) | -1.652 (4.775) | -1.424 (5.147) | 1.037 (4.663) | .555 (4.534) | .053 (5.847) | -1.334 (4.887) |
| Total patents | -1.349* (.760) | .093 (1.596) | -.840 (1.401) | -.034 (1.508) | -1.487* (.821) | -.286 (1.546) | -.532 (1.220) | 1.089 (1.479) |
| Patent quality | 3.149*** (.913) | 3.075*** (.950) | 2.234* (1.149) | 2.278** (1.070) | 4.370*** (1.101) | 3.777*** (1.141) | 3.049** (1.280) | 2.581* (1.314) |
| Year dummies | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 363 | 363 | 363 | 363 | 363 | 363 | 363 | 363 |
| Over-identification, p-value | .309 | .488 | .764 | .776 | .159 | .307 | .505 | .342 |

Notes: 1) The full set of IVs include IBM's opposed patents and IBM's opposed patents X afteryear2003. 2) Robust standard errors, clustered by market, are in parentheses. 3) * significant at 10%, ** significant at 5%, *** significant at 1%. 4) We use Wooldridge's quasi-differencing transformation to remove market fixed effects.