Competition Between Open Source and Proprietary Software: Strategies for Survival

Michael Sacks*
Department of Economics
University of California, Irvine
3151 Social Science Plaza A
Irvine, CA 92697
msacks@uci.edu

BIO: Michael Sacks is a Ph.D. candidate in the department of economics at the University of California, Irvine. He holds an M.A. in mathematical behavioral sciences. His research interests are in the economics of software development and in the dynamics of collaborative production.

ABSTRACT: There are two puzzles in the software competition literature: will both proprietary and open source software survive and how do producers of proprietary software differentiate themselves from open source competition. I address both puzzles by analyzing competition between a firm producing proprietary software and a community producing open source software. If the firm faces no competition, then the software caters to less technologically savvy individuals. When facing competition, the open source software caters to the most technologically savvy individuals, leading the firm to target even less savvy individuals than it would when acting as a monopolist in order to differentiate its software from the open source option. The open source movement, then, may not be an unalloyed success as the growth in open source can be tied to deterioration in the proprietary software. Given that both types of software survive by catering to different segments of the market, an important avenue for research will be to analyze the stability of the underlying segments and the corresponding welfare implications.

KEYWORDS: Hotelling, Mixed Duopoly, Differentiation, Heterogeneous Consumers, Endogenous Fixed Costs, Open Source, Proprietary, Software.

* Thanks to Jean-Paul Carvalho, Jiawei Chen, Linda Cohen, Michael McBride, and two anonymous referees for helpful comments. All errors are my own.

Introduction

The software industry is typically characterized by two distinct development methodologies: proprietary software (PS) and open source software (OSS).[1] By its nature, OSS is ``free'' in that the source code (human readable code) is made publicly available at no cost, so any individual can conceivably download the source code, compile it into binary code (machine readable code), and run it on her computer. To the contrary, the only component made available by the PS developer is the binary code. In this paper, I address a fundamental question concerning competition: which development methodology will survive in the long run?

Popular media suggests that OSS is the future [40,44]. I argue that which development methodology represents the future is irrelevant. The developers of PS and the developers of OSS need not compete over the same market. There are, of course, overlaps between the target markets; however, the underlying nature of the OSS development process leads to self-selection, where the PS developer differentiates its product to focus on the market not targeted by the OSS. Therefore one should expect both development methodologies to survive in the long run, so long as the underlying markets the two methodologies target continue to exist.

Consider the market for scientific typesetting software, specifically two products: the proprietary Scientific Word (SW) and the open source GNU Emacs (Emacs). SW is a "what you see is what you get (WYSIWYG)" editor, akin to Microsoft Word, while Emacs is a "what you see is what you mean (WYSIWYM)" editor, which more closely resembles

[1] More recently, a third methodology - mixed source - has been identified as a hybrid between open source and proprietary (closed source). See [11] for details.

computer programming to document typesetting. SW is designed for those who routinely typeset documents heavy in mathematical notation and collaborate with others who do the same, such as LaTeX users. Emacs is fully capable of replicating every feature of SW, but can also be used in other applications, including computer programming and editing system files. Furthermore, Emacs is known for its immense customizability and steep learning curve.

Thus two more appropriate questions to ask are (i) who does each type of developer target, and (ii) are there externalities to competition between the two processes? In this paper, I address the first question. There is a large literature involving network effects and positive externalities in open R&D (R&D without, or with minimal intellectual property restrictions) addressing the second question, including work directly on software competition, such as [37], [17,18], [48], [29], [13], [33], and [4]. For a review of the theory of network industries, see [49] and [50]. A review of the empirical literature is provided by [5].

Before researchers began analyzing OSS, many of the features of software competition were captured by the standard models of quality competition and durable goods markets [8,9,51,52]. Once OSS came into fruition, researchers began developing new models of competition. Two primary classes of models emerged from this research program. The first class analyzes externalities and spillover effects and the second class determines long-run shares and survival. The second class has received considerably less attention than the first. See [28] for a review of these avenues of research in OSS. A brief survey of the economics of OSS is provided by [22].

[36] and [12] develop general results related to long run market shares in industries with network effects and show there exist conditions under which multiple firms survive., [7], [10], [45], [48], [26], [13], [15], and [33] were among the first to analyze the long-run market shares in the software industry when there is competition between PS and OSS. Each paper arrives at a similar conclusion: it depends on either initial conditions or exogenous parameters. That is, there are scenarios in which OSS overtakes the market, those in which PS is the lone survivor, and those with coexistence.

[48] models software usability when there is competition between PS, OSS, and a commercial OSS developer, where consumers vary according to their preference for usability, but have an identical inherent value of the software, regardless of the developer. The PS (OSS) is assumed to be the most (least) usable software. Only the commercial OSS developer is strategic in its usability decision, and must locate somewhere in between the PS and the OSS. The author finds that, if network effects are weak, then the commercial OSS developer has the incentive to make its software less usable than the PS. If network effects are strong, then the commercial OSS developer develops software as usable as the PS, which drives the PS developer out of the market. However, this equilibrium emerges due to two factors: the consumers value all software identically, and the usability of the PS is exogenously set to the minimal value. If these assumptions are dropped, and both the PS and commercial OSS producers each endogenously select their usability, then multiple equilibria emerge and the conclusions no longer apply.

[26] provide an elegant analysis of competition between OSS and PS, focusing on vertical quality and pricing decisions of the PS developer. The authors find that quality is inversely related to competition from OSS; that is, the presence of OSS decreases the

quality of PS, regardless of the initial quality of the OSS. The effect is intensified as the quality of OSS is increased. On the other hand, when facing competition from other PS developers, quality can go in any direction depending on how closely substitutable the programs are. Consumers are assumed similar in their valuations of quality. That is, an increase in quality necessarily leads to an increase in demand. It is reasonable to consider quality to have a horizontal component to it, where different consumers have different needs to be satisfied, and quality is determined by how well those needs are met. Under a horizontal valuation such as the one described above, the same comparative statics can break down. For instance, an increase in the quality of OSS can in fact lead to an increase in the quality of PS, depending on the distribution of consumers.

In this paper, I treat the degree of product differentiation as a strategic choice and analyze consumer choice through this endogenous process, developing a model in which there is a single firm developing PS and a potential consumer base. The firm sequentially chooses how sophisticated the software should be via an investment in technology, and then sets the price in order to maximize profits. Competition arises through the development of OSS, which is developed by a subset of the potential consumer base. The community producing the OSS chooses the level of sophistication of the OSS through a similar investment in technology at the same time as the firm. Rather than maximizing profits, the community seeks to cooperatively find an efficient level of sophistication of the OSS.[2] Each individual has a preferred level of technological sophistication. Individuals then choose which option to purchase and face two costs: the price of the software (which

[2] Equivalently, a sequential non-cooperative bargaining process in which individuals seek to maximize their own utility generates an equivalent outcome if the individuals are sufficiently forward looking.

is zero for the OSS) and a learning cost, which is determined by the consumers' preferred level of technology and the actual employed level of technology.

Competition is modeled as a mixed duopoly, though this model differs from the standard mixed duopoly.[3] Typically, mixed duopolies are modeled with a private, profit-maximizing firm competing with a government-owned entity, concerned with maximizing social welfare. More recently, researchers have developed models of partial-privatization, where a fraction of the firm is owned by individual shareholders concerned with maximizing profits while the remaining shares are held by a welfare-maximizing government [3,34]. In this paper, the OSS developers are interesting in maximizing their own welfare. It is similar to a model of privatization except that there are distinct asymmetries between the objective functions of the OSS developers and the firm.

I find that, when there is competition between a firm developing PS and a community developing OSS, the firm will cater the software towards individuals with lower technological ability, a phenomenon I call *catering to ignorance*. Furthermore, this unique equilibrium emerges endogenously. This notion is not new. [32] reference a quote from an OSS developer, who states:

> [I]n every release cycle Microsoft always listens to its *most ignorant customers*. This is the key to dumbing down each release cycle of software for further assaulting the non personal-computing population. Linux and OS/2 developers, on the other hand, tend to listen to their *smartest* customers… The good that Microsoft does in bringing computers to non-users is outdone by the curse that they bring on experienced users [39].

[3] See [23] and citations therein for a survey of the early models of mixed oligopoly.

[32] illustrate the importance of the argument made in the above quote to license selection. To my knowledge, however, this is the first paper to provide an analytical microeconomic foundation leading to this endogenous differentiation under consumer heterogeneity. [27] alludes to this idea in a brief extension by considering an environment with modular (piecewise) development, where as the number of modules grows, the probability of OSS developing all modules converges to zero while the probability converges to a positive value for the PS.

The OSS applies downward pressure, i.e., the firm's choice of software sophistication decreases relative to a monopolist. In reference to the above quote, it can be said that Microsoft listens to its least technological savvy consumer base *because* Linux developers listen to the most technologically savvy individuals, namely themselves.

The endogenous equilibrium emerges from either one of two possible mechanisms: heterogeneous costs of technology and the intrinsic motivations of OSS contributors. Much of the work on software competition has ignored the fact that technology is not easy to use, or equivalently, it is costly (an exception being [16]). Therefore the real price of using software consists of two components, the purchase price and a learning cost. The learning cost is not the same for all individuals. Some individuals are more technologically savvy and are able to learn to use technology at a lower cost, in terms of time and effort. Thus any two individuals purchasing the same software likely face different prices (in real terms), which even compels a monopolist PS developer to choose a degree of technology below the median ability of the consumer population.

Even without heterogeneous learning costs, this same pattern can emerge. The intrinsic motivations of contributors to OSS tends to bias the chosen technology level of the software above the median, which limits the market available to the firm, since it needs to compete with a zero-price competitor to capture high-skill individuals. In response, the firm best responds by targeting the group most ignored by the OSS, the technologically ignorant individuals.

The Open Source Software Paradigm

This section is divided into two parts. The first defines OSS while the second expounds upon the process of its development.

Open Source Software: Definition and Background

The Open Source Initiative (OSI), a non-profit 501(c)3, was founded in 1998 to formalize the open source definition (OSD). OSI is the community-recognized body for reviewing and approving licenses as conforming to OSD.[4] Those licenses furthest from the standard copyright licenses are typically referred to as copyleft. Among the most copyleft license is the GNU General Purpose License (GPL) and among the most copyright, but still considered conforming to the OSD is the Berkeley Software Distribution 2-Clause (BSD) [41,42].

The implication is simple: free software is not equivalent to OSS. As a consumer, there is still a cost of consuming OSS. In its entirety, the source code principle in the OSD states:

> The program must include source code, and must allow distribution in source code as well as compiled form. Where some form of a product is not distributed with source code, there must be a well-publicized means of obtaining the source code for

[4] http://opensource.org/about, accessed January 27, 2015.

> no more than a reasonable reproduction cost preferably, downloading via the Internet without charge. The source code must be the preferred form in which a programmer would modify the program. Deliberately obfuscated source code is not allowed. Intermediate forms such as the output of a preprocessor or translator are not allowed [43].

At a minimum, the source code must be made available; however, computers cannot directly read source code. It must be compiled into binary, which computers are then able to interpret. This fact implies that, even in instances where the source code is made freely available, the software is not free in the economic sense. It is costly in terms of time and knowledge. The same holds true in PS, where users must invest time and energy in learning to use the software. To capture this feature, I treat all software as costly, where the cost is characterized by the sophistication of the software's technology. The cost of learning to use the software depends on how technologically skilled the user is and how technologically sophisticated the software is. For the purposes of this paper, I assume that consumers are able to freely obtain pre-compiled binaries of OSS. Thus individuals face both a price and learning cost when purchasing the PS and only a learning cost when using the OSS. The environment can easily be extended to one in which consumers must expend extra effort to compile the source code; however, this would only make the product differentiation more extreme, as the PS becomes more valuable to the marginal consumer, relative to the OSS.

Open Source Software: Development

This section's focus is on the primary characteristics relevant to this paper, namely who contributes, how development decisions are made, and why individuals contribute [1,2,6].

In general, there are three kinds of contributors: institutional contributors, individuals representing institutions, and hobbyists. Institutional contributors can further be subdivided into two types: professional OSS entities and commercial supporters. Note that these types need not be mutually exclusive. For example, Red Hat Inc. provides OSS and, for a fee, related services to enterprises. In this respect, it is acting as a professional OSS entity. However, Red Hat also sponsors the Fedora Project, where it is acting as a commercial supporter.[5] While Fedora is primarily a community-driven endeavor, Red Hat has veto power over all Fedora development decisions.

Institutional contributors typically support large-scale OSS projects. A notable example is Mozilla's OSS (such as Thunderbird and Firefox). The Mozilla Suite is made up of all types of contributors. There are three subsidiaries to consider: the Mozilla community, the Mozilla Corporation, and the Mozilla Foundation. The Mozilla community consists of over 10,000 contributors.[6] The Mozilla Corporation is a taxable entity that serves its non-profit, public parent (the Mozilla Foundation). Its leadership, known as the ``Steering Committee'' is responsible for guiding the direction of the projects. In short, the Mozilla Corporation works to fund the Mozilla Foundation, which guides the direction of the projects developed by the Mozilla community.

The preceding two examples illustrate the underlying nature of how development decisions are made. Typically collaboration occurs online through email lists and message boards through hosting servers, such as Github or Sourceforge. The project's leadership facilitates communication back and forth between various contributors and establishes a plan of action is established according to the input of contributors. Rather than explicitly

[5] The Fedora Project coordinates the development of Fedora, a Linux-based operating system.
[6] https://www.mozilla.org/en-US/contribute/, accessed January 28, 2015.

modeling this bargaining process, I employ the results of the median voter theorem, as preferences are assumed to be single peaked.

[31] identify several theoretical reasons why individuals may contribute to OSS projects without remuneration. The reasons can be classified into two categories: immediate rewards and delayed benefits. Immediate rewards include productivity gains, enjoyment from working on the project, and peer recognition, while delayed rewards include various signaling effects and future employment opportunities. These effects are expounded upon in various works, such as [32], [30], [35], [6], [21], [47], and [20]. Based on a survey conducted by [24], some programmers claimed that signaling was a motivating factor in their decision to contribute. According to [25], individuals expend more effort signaling when performance is more visible to the firms, performance has a greater impact, and performance is highly informative about skill. [38] develop an evolutionary model explaining voluntary contributions to OSS. To capture these features, I model the benefit of contributing to OSS as increasing in the technological sophistication of both the programmer and the software

The Model

There is a single firm developing PS utilizing technology level $t \in [0,1]$ and charging price $p \geq 0$. The firm is not alone in the market. It faces competition from OSS, which utilizes technology level $\tau \in [0,1]$ and is sold at a price of zero. I refer to an arbitrary investment by ζ. The firm does not face a development cost; however, the developers of OSS do face one, to be specified shortly. This assumption on the cost function of the firm is fairly innocuous, and I will show that the results are qualitatively unchanged for both a cost function increasing in t and a U-shaped cost function.

There is a continuum of consumers normalized to unit mass. Each consumer is identified by her technological prowess $x \in [0,1]$. Technological prowess distributed amongst the consumers according to the distribution $G(\cdot)$ with associated log-concave pdf $g(\cdot)$. Individuals with $x \leq G^{-1}\left(\frac{1}{2}\right)$ are said to be *technologically ignorant*, while individuals with values $x \geq G^{-1}\left(\frac{1}{2}\right)$ are considered to be *technologically elite*. Every consumer faces unit demand for software with single-peaked valuation $V = v(x, \zeta)$, and chooses either the PS, the OSS, or an outside option normalized to zero value. V is twice differentiable in each argument, symmetric in ζ and x, strictly concave in x and ζ, and obtains a maximum at $v(x, x) = \bar{V}$ at any x. Each individual has an ideal technology level suited for her own technological prowess, and obtains equal loss when the software's location varies from this value, regardless of the side of x that ζ falls on.[7] The symmetry assumption does not imply that utility itself is symmetric. In fact, utility is asymmetric due to the increasing learning cost, to be described.

Consumers also face a cost of learning to use the software, which depends on their own technological prowess x, and the level of technology employed by their chosen software ζ, as illustrated in [16]. Define this cost as $\ell(x, \zeta)$, which is assumed twice differentiable in each argument, decreasing and weakly convex in x and both increasing and strictly convex in ζ. At $x = 1$, $\ell(1, \zeta) = 0$ for all ζ and at $\zeta = 0$, $\ell(x, 0) = 0$ for all x.

[7] Symmetry allows a wide variety of functions, such as the power loss function $v(x, \zeta) = \bar{V} - \theta(x - \zeta)^{\alpha}$, $0 < \theta \leq \bar{V}$, $\alpha = 2k, \forall k \in \mathbb{N}$.

Figure 1. Approximately Here

That is, the most technologically elite individual can learn even the most sophisticated software at no cost, while any individual can learn to use the least sophisticated software at no cost. Figure 1 graphically illustrates this relationship: panel (a) with respect to the technological investment and panel (b) with respect to individual prowess. Any individual can use any level of software at a cost. For any given level ζ, the cost of learning to use the software is decreasing in x. Thus a technologically elite individual can learn to use a technologically advanced software at a lower cost than a technologically ignorant individual, and can also learn to use a low technology software at a lower cost than a technologically ignorant individual.

OSS is developed by a subset of the individuals in the economy. Contributors receive two types of benefits: a benefit from using the software $V = v(x,\tau)$, and a benefit from contributing. Define the contribution payoff as $b(x,\tau)$, which is assumed to be twice differentiable in each argument, nondecreasing and weakly concave in x and increasing and weakly concave in τ. The cost of contributing is defined by the function $K(x,\tau)$, where $\frac{\partial}{\partial x}K(x,\tau) < 0$, $\frac{\partial^2}{\partial x^2}K(x,\tau) > 0$, and $\frac{\partial}{\partial \tau}K(x,\tau) > 0$. While variables such as time/effort do not enter the cost/benefit function, under the rational choice framework employed in this paper, any individual with ability x contributing to software τ would choose an identical effort level, as the value which equates marginal benefit and marginal cost would be identical. Thus the model can be seen as isomorphic to one in which time effort is explicitly modeled, with the output/notation suppressed. Furthermore, I assume that

$$b(1,\tau) > K(1,\tau), \forall \tau. \tag{1}$$

This condition guarantees that a nonzero measure of individuals are willing to contribute to develop the OSS.

Utility is assumed quasilinear, which implies that the utility function of an individual x can be written as

$$u(x,t,\tau,p)=\begin{cases} v(x,t)-\ell(x,t)-p & \text{if consuming PS} \\ v(x,t)+b(x,\tau)-\ell(x,t)-K(x,\tau) & \text{if consuming PS \& producing OSS} \\ v(x,\tau)+b(x,\tau)-\ell(x,\tau)-K(x,\tau) & \text{if consuming \& producing OSS} \\ v(x,\tau)-\ell(x,\tau) & \text{if only consuming OSS} \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

It also follows that $u(x,t,\tau,p)$ is strictly concave and single peaked in x.

Outline of the Game

Interactions occur over the course of three stages. In the first stage, the firm and OSS contributors simultaneously and independently select t and τ, respectively. The firm's objective is to maximize profits and the OSS developers' objective is to maximize the total utility of contributors. In the second stage, the firm chooses price $p(t)$ given the technology investment and the OSS price is fixed at zero. In the third stage, individuals make their purchase/usage decision. Since information is complete, I will be looking for a subgame-perfect Nash equilibrium.

Equilibrium Analysis

The equilibrium analysis consists of two parts. I first analyze a monopolist firm developing PS, choosing both location and price, to establish a benchmark to characterize the preferred firm location. I then add an OSS and compare the firm's decision under each scenario.

Throughout the analysis, I will be using three terms: the lower neighborhood about x, defined as $N_\eta^-(x) \equiv \{y\colon 0 \leq x - y \leq \eta\}$, the upper neighborhood about x, defined as $N_\eta^+(x) \equiv \{y\colon 0 \leq y - x \leq \eta\}$, and the neighborhood about x, defined as $N_\eta(x) \equiv N_\eta^-(x) \cup N_\eta^+(x)$.

Monopoly Benchmark with only Proprietary Software

Consider the third stage of the game. For a given equilibrium price and location $(p(t^m), t^m)$, consumers will purchase the PS if

$$v(x, t^m) - p(t^m) - \ell(x, t^m) \geq 0. \quad (3)$$

Taking (3) into account, the firm chooses the price p. The following lemma characterizes the demand function faced by the firm.

Lemma 1 *For any given technology investment t^m, the support of the demand function of the firm, is a non-degenerate convex subset of* $[0,1]$.

Lemma 1 implies that there are at most two critical individuals.[8] Order these individuals by their technological prowess x and define the lower-critical individual by $\underline{x}(p, t^m)$ and the upper-critical individual by $\overline{x}(p, t^m)$. The demand function takes the form

$$D(p, t^m) = \int_{\underline{x}(p,t^m)}^{\overline{x}(p,t^m)} g(x)dx, \quad (4)$$

where $0 \leq \underline{x}(p, t^m) \leq \overline{x}(p, t^m) \leq 1$. The firm can either target the entire market, an interior subset $[\underline{x}(p, t^m), \overline{x}(p, t^m)]$, or a subset including the boundary; either $[0, \underline{x}(p, t^m)]$ or $[\overline{x}(p, t^m), 1]$, with $\underline{x}(p, t^m) > 0$ and $\overline{x}(p, t^m) < 1$.

[8] When interior, these individuals are indifferent between purchasing the PS and choosing the outside option.

The next step in describing the demand function is to define properties of $\underline{x}(p, t^m)$ and $\overline{x}(p, t^m)$. Lemma 2 does this and provides a relevant implication for the demand function.

Lemma 2 $\frac{\partial \underline{x}(p,t^m)}{\partial p} \geq 0$ *and* $\frac{\partial \overline{x}(p,t^m)}{\partial p} \leq 0$. *If* $0 < D(p, t^m) < 1$, *then at least one of the inequalities must hold strictly, which implies that* $\frac{\partial D(p,t^m)}{\partial p} \leq 0$.

Together, lemmas 1 and 2 show that demand is downward sloping. Furthermore, there is a region in which the demand is inelastic and a region in which demand is elastic. In the second stage, the firm's objective is to choose the price that maximizes profits, for every possible technology investment t.

Proposition 1 *In equilibrium, the pair* $(p(t^m), t^m)$ *must satisfy* $\varepsilon\big(p(t^m)\big) = -1$, *where* $\varepsilon\big(p(t^m)\big)$ *is the demand elasticity at price* $p(t^m)$.

This is the standard profit maximizing result. A firm with zero marginal cost of production sets the price such that demand elasticity equals unity.[9]

Peeling the game tree back to the first stage, the firm optimally selects $t = t^m$, understanding that it will then choose price $p(t^m)$. That is, price is treated as a function of location. This is in contrast to [48], where the location t^m is assumed to equal zero non-strategically.

Proposition 2 *The equilibrium technological investment is interior and satisfies* $\mu(t^m)$, *where* $\mu(t^m)$ *is the elasticity of technology, defined by* $t^m \frac{\frac{\partial D(p(t^m),t^m)}{\partial t}}{D(p(t^m),t^m)} = 0$.

[9] If there was a positive marginal cost of production, then proposition 1 would require $\varepsilon\big(p(t^m)\big) < 1$, as is standard.

The choice of t^m affects demand through two channels: a demand (direct) effect through t^m and a strategic (indirect) effect through $p(t^m)$. Given the sequential nature of the problem, the firm need only consider the demand effect since the price will self adjust in the second stage. Thus the objective of the firm collapses to choosing the t^m, which conditional on $p(t^m)$ maximizes the firm's market share.

The intuition that t^m is interior is straightforward. Suppose that the firm chooses location $t^m = 0$ and price $p(0)$. At this location and price, the individual with the highest valuation (and thus willingness to pay) is the one located at $x = 0$. Now consider the upper neighborhood about $x = 0$. These individuals must also have a strictly positive valuation. It follows that the monopolist could increase the technology level to $t^m = 0 + \xi$ for some small, positive ξ and increase the size of its market while still charging a price of $p(0)$, thereby increasing profits since $\underline{x}(p(0), 0) = \underline{x}(p(0), \xi)$ and $\overline{x}(p(0), 0) < \overline{x}(p(0), \xi)$ for small enough ξ.[10] A similar argument holds at $t^m = 1$ and $p(1)$.

It immediately follows that the individuals located at $x = 0$ and $x = 1$ should never obtain positive surplus. If the individual at $x = 1$ obtains a positive surplus at technology level t, then it must be that $x = 1$ has positive surplus at $t - \xi$ for small enough ξ. Thus, not only are $t = 0$ and $t = 1$ ruled out, but also all values of t arbitrarily close to zero and one. Furthermore, since $g(\cdot)$ is log-concave, there is higher mass placed at values of x near the median than the extremes.[11] Thus $G^{-1}\left(\frac{1}{2}\right) \in \left[\underline{x}(p(t^m), t^m), \overline{x}(p(t^m), t^m)\right]$ must be true. By continuing the analysis, the following result emerges.

[10] This line of reasoning assumes that the market share is currently a strict subset of [0,1]. A similar argument can be made when the firm's market share covers the entire consumer base.

[11] The exception is the uniform distribution, which satisfies log-concavity and places equal mass across all x in the support.

Figure 2. Approximately Here

Proposition 3 *In equilibrium,* $t^m < G^{-1}\left(\frac{1}{2}\right)$.

Compared to the typical Hotelling model, a bias is introduced through the learning cost, which pushes t^m to the left of the median, towards technological ignorance. Therefore the OSS developer quoted by [32] was correct in saying that the firm "targets" the technologically ignorant consumers.

Thus the subgame-perfect equilibrium outcome can be characterized as follows. The firm chooses location $t^m < G^{-1}\left(\frac{1}{2}\right)$ and price $p(t^m)$ such that $p(t^m)$ and t^m satisfies the conditions in propositions 1 and 2. All consumers $x \in \left[\underline{x}(p(t^m), t^m), \overline{x}(p(t^m), t^m)\right]$ purchase the PS. The remaining individuals choose the outside option. Figure 2 provides a unidimensional illustration of the equilibrium.

Mixed Duopoly

The analysis again proceeds through a standard application of backward induction. Taking equilibrium locations t^d, τ^d and price $p(t^d, \tau^d)$ as given, there are six possible combinations: an individual can choose to purchase the PS, with or without contributing to the OSS, use the OSS without contributing, both contribute to and use the OSS, only contribute to the OSS and use the outside option, or just choose the outside option.

In equilibrium, an individual with technological prowess x prefers to purchase the PS if

$$v(x, t^d) - p(t^d, \tau^d) - \ell(x, t^d) \geq \max\{v(x, \tau^d) - \ell(x, \tau^d), 0\}, \quad (5)$$

which, as in the monopolist case, is linear in price. Note that $b(x, \tau^d)$ and $K(x, \tau^d)$ are not included. This omission is because the contribution decision is made in a previous stage, and therefore both the benefit and cost are sunk. Consider the critical individuals, characterized by technological prowesses $\underline{x}(p(t^d,\tau^d),t^d,\tau^d) \geq 0$ and $\overline{x}(p(t^d,\tau^d),t^d,\tau^d) \leq 1$. Lemmas 1 and 2 apply in this case as well, so $\underline{x}(p(t^d,\tau^d),t^d,\tau^d)$ and $\overline{x}(p(t^d,\tau^d),t^d,\tau^d)$ have the same properties as in the monopoly case.

Given these critical consumers, the demand function for the firm, given equilibrium technology levels t^d and τ^d, is

$$D(p,t^d,\tau^d) = \int_{\underline{x}(p,t^d,\tau^d)}^{\overline{x}(p,t^d,\tau^d)} g(x)dx. \tag{6}$$

The associated profit function is $\pi(p,t^d,\tau^d) = pD(p,t^d,\tau^d)$. Profits are zero when either price is zero or demand is zero, so order for the firm to obtain positive profits, it must be that $p(t^d,\tau^d) > 0$ and $D(p(t^d,\tau^d),t^d,\tau^d) > 0$. That is, the firm is profitable only when it charges a positive price and it faces positive demand. The following proposition begins to characterize the conditions under which this statement is true.

Proposition 4 *In any equilibrium where* $p(t^d,\tau^d) > 0$*, the firm faces positive demand only if* $t^d \neq \tau^d$.

In other words, product differentiation endogenously arises through competition. Suppose a rational consumer is given a choice between two identical options, where one is offered at a positive price, while the other is offered at zero price. The consumer always chooses the zero price option. The firm could set $t^d = \tau^d$ and $p = 0$ in an effort to split the market with the OSS developers, but it always has a more profitable option via product differentiation.

Proposition 4 shows that all consumers who do not choose the outside option choose the OSS when there is no product differentiation, so the firm has the incentive to distinguish its PS from the OSS. The next logical question is when there is product differentiation, how is the market segmented? I proceed by jointly analyzing the two subgames induced by the technology choices of the firm and the OSS developers, $t^d < \tau^d$ and $t^d > \tau^d$.

First suppose that $t^d < \tau^d$. It is obvious that $\overline{x}(p, t^d, \tau^d) < \tau^d$ since any individual in the neighborhood of $x = \tau^d$ consumes open OSS. Now suppose that $t^d > \tau^d$. In this case, $\underline{x}(p, t^d, \tau^d) > \tau^d$ since the individuals located in the neighborhood of $x = \tau^d$ consumes the OSS. The following proposition characterizes the price under either of the two cases.

Proposition 5 *The equilibrium price* $p(t^d, \tau^d)$ *must satisfy* $\varepsilon\left(p(t^d, \tau^d)\right) = -1$ *for all feasible* t^d, τ^d*.*

Proposition 5 is analogous to proposition 1.

The final two stages of the game can be summarized as follows:

(i) In stage 3, all consumers in the interval $\left[\underline{x}(p(t^d, \tau^d), t^d, \tau^d), \overline{x}(p(t^d, \tau^d), t^d, \tau^d)\right]$ purchase the PS. If $t^d < \tau^d$, then every individual located on the interval $[0, \underline{x}(p(t^d, \tau^d), t^d, \tau^d))$ chooses the outside option and every individual located on the interval $(\overline{x}(p(t^d, \tau^d), t^d, \tau^d), 1]$ chooses the OSS. If $t^d > \tau^d$, then every individual located on the interval $[0, \underline{x}(p(t^d, \tau^d), t^d, \tau^d))$ chooses the OSS and every individual located on the interval $(\overline{x}(p(t^d, \tau^d), t^d, \tau^d), 1]$ chooses the outside option.

(ii) In stage 2, the firm sets the price $p(t^d, \tau^d)$ according to proposition 5.[12]

[12] Define $[0,0) = (1,1] = \emptyset$, i.e., no individual consumes the associated software.

Focusing on the first stage, the firm and the OSS developers simultaneously choose t^d, τ^d to maximize their respective objective functions. At the extensive margin, a consumer x will contribute to the OSS if

$$b(x, \tau^d) \geq K(x, \tau^d) \tag{7}$$

Let $\hat{x}(\tau^d)$ represent the value of x such that (7) holds with equality. Every individual $x \geq \hat{x}(\tau^d)$ will contribute to the OSS. The remaining individuals do not.

The development of OSS is a cooperative social process, so the collaboration that leads to the choice of location for the OSS can thus be described by a social choice function. Define $X(t) = \{x | x \geq \hat{x}(t)\}$ as the set of all individuals who contribute to the OSS. The social choice function $F: X(t) \to [0,1]$ prescribes a location τ^d based on the individuals $x \in X(t)$. A reasonable strategy proof social choice function is the smallest median peak.

While the imposition of this social choice function may seem *ad-hoc*, [14] show that in a non-cooperative bargaining problem, the set of proposals that can be passed in any pure strategy subgame-perfect equilibrium collapses to the ideal point of the median individual, conditional on a sufficiently low discount rate. The social choice function can be interpreted as either the level of sophistication that maximizes total contributor welfare, or equivalently as the outcome of an unmodeled, game theoretic bargaining process in which each individual seeks to maximize her own utility. In either case, the OSS developers' choice τ^d satisfies the following two equations:

$$G(\tau^d) - \frac{1}{2} G\left(\hat{x}(\tau^d)\right) = \frac{1}{2} \tag{8}$$

$$b(\hat{x}(\tau^d), \tau^d) = K(\hat{x}(\tau^d), \tau^d). \tag{9}$$

Figure 3. Approximately Here

Equation (8) pins down the conditional median, conditional on the indifferent individual $\hat{x}(\tau^d)$ and equation (9) pins down the indifferent individual. It remains to be shown that such a value $\hat{x}(\tau^d)$ exists.

Lemma 3 *For every t, there exists a value $\hat{x}(t)$ such that $b(\hat{x}(t), t) = K(\hat{x}(t), t)$.*

The benefit of contributing is increasing in x while the cost of contributing is decreasing in x. Figure 3 illustrates the indifference point.

Note that $\tau^d \geq G^{-1}\left(\frac{1}{2}\right)$ since the conditional median cannot be lower than the unconditional median given the cost of contributing $K(x, t)$. Choosing $\tau^d \geq G^{-1}\left(\frac{1}{2}\right)$ is what I call *catering to elites*, where the OSS is designed specifically with the technologically elite individuals in mind. Also the OSS developers' optimal choice is independent of the firm's choice t^d. That is, no consideration is given to any individual who is not a participant in the development process, implying that the firm's problem can be reduced to maximizing profits, conditional on τ^d, which simplifies the analysis since expectations do not need to be taken over the distribution of the conditional median. As in the monopolist case, the firm's equilibrium location choice must satisfy

$$\frac{\partial D(p(t^d, \tau^d), t^d, \tau^d)}{\partial t} = 0$$

if interior.

The important question is where does the firm locate, relative to the OSS? It turns out that it is always more profitable for the firm to locate to the left of the OSS. That is,

Figure 4. Approximately Here

$t^d < \tau^d$. I call this relationship *catering to ignorance*, where the firm targets the technologically ignorant individuals while the OSS is designed for the technologically elite.

Proposition 6 *In equilibrium, the firm locates to the left of the OSS developers:* $t^d < G^{-1}\left(\frac{1}{2}\right) < \tau^d$.

Since $\tau^d \geq G^{-1}\left(\frac{1}{2}\right)$, the market share of the firm at any given price is at most $\frac{1}{2}$ if the firm locates at a point such that $t^d \geq \tau^d$. By moving just slightly to the left of τ^d, the firm weakly increases its market share without changing its price. In summary, given that the OSS caters to the elite, the firm is strictly better off catering to ignorance.

The next step is to examine how t^d and t^m relate.

Proposition 7 *The equilibrium technological sophistication of the PS in the mixed duopoly is weakly lower than the location of the PS in the monopoly.*

Demand for the PS is constrained from the right under the presence of OSS. Therefore, either one of two events must be true if the PS under the duopoly is located to the right of the PS under the monopoly. The monopolist firm would always have the incentive to replicate the duopoly outcome if the duopoly firm is doing better, or vice versa. The catering to ignorance result is thus twofold. First, the PS is developed with the technologically ignorant individuals in mind while the OSS is developed with the technologically elite individuals in mind. Second, the presence of OSS actually pushes the PS further toward the most technologically ignorant individuals. Figure 4 provides a unidimensional representation of the subgame-perfect equilibrium.

The easiest market to notice this location choice is in the market for computer operating systems. Consider Mac OS X. At its current stage, OS X is relatively easy to learn, use, and maintain. Compare this (proprietary) option to an open source operating system, such as Arch Linux. Installing OS X is as simple as pushing an install button. To install Arch Linux, first the user must download and verify the install medium, then change the boot order so that the install medium (typically a USB drive or CD) is read before the hard disk drive or solid-state drive. Once the medium is opened, there is no graphical user interface (GUI); all commands are entered in an Unix-like environment. Even when the install is finished, the user must manually install a GUI. The difficulty of use for a technologically ignorant user does not end with the installation. System maintenance and upgrades must also be conducted manually in an Unix-like text environment. If the user is not careful, upgrading can damage the system when certain incompatible packages are combined and dependencies are broken.

Lastly, I analyze how changes in $\hat{x}(\tau^d)$ impact the firm's choice t^d.

Proposition 8 t^d *is weakly increasing in* τ^d. *The inequality is strict if* $t^d < t^m$ *and* $t^d > 0$.

Proposition 8 is an implication of the firm's desire to profit as if it were a monopolist. If the OSS locates itself towards more sophisticated individuals, then those individuals to the left of the lower neighborhood about $x = \tau^d$ are more willing to purchase the PS, allowing the duopolist firm to capture a market that more closely resembles its monopolist counterpart. This result is in contrast to [26], who find an inverse relationship between the quality (location) of PS and OSS, as opposed to the positive relationship this paper finds. The difference can be attributed to how heterogeneity is modeled. [26] assume purely vertical differentiation, while I allow for both horizontal and vertical differentiation through v and ℓ.

Consider the evolution of Microsoft Word over time, namely with typesetting mathematical equations. Prior to Word 2013, the Equation Editor was required to add mathematical equations to documents. As open source LaTeX options have become easier to use, Microsoft has responded by including a built in equation menu, where equations can be added with the push of a button.

Analytical Example

To better illustrate the results, consider the following example. Suppose that

$$G \sim U(0,1)$$

$$v(x,\zeta) = 1 - (x-\zeta)^2$$

$$\ell(x,\zeta) = (1-x)\zeta^2.$$

For expositional convenience, I will suppress the monopoly and duopoly equilibrium superscripts m and d when there is no confusion.

If the firm holds a monopoly, then in the third stage, consumer x will purchase the PS if

$$1 - (x-t)^2 - p - (1-x)t^2 \geq 0$$

Since G is uniform, $D(p,t) = \overline{x}(p,t) - \underline{x}(p,t)$. Solving the above inequality for x yields the associated demand function

$$\min\left\{1, \frac{1}{2}\left(2t + t^2 + \sqrt{4 - 4p - 4t^2 + 4t^3 + t^4}\right)\right\}$$

$$-\max\left\{0, \frac{1}{2}\left(2t + t^2 - \sqrt{4 - 4p - 4t^2 + 4t^3 + t^4}\right)\right\},$$

which is clearly weakly decreasing in price as required. The first line represents $\overline{x}(p,t)$ and the second line represents $\underline{x}(p,t)$.

Fixing $t = t^m$ and using standard optimization techniques, it is straightforward to show that the equilibrium strategy is to choose p such that

$$p(t^m) = 1 - 2(t^m)^2,$$

$$t^m = \frac{1}{1+\sqrt{2}} < \frac{1}{2}.$$

The median of $G(x)$ is $\frac{1}{2}$, so t^m is strictly below the median as expected. The reason is the learning cost $\ell(x,t)$. Suppose that $\ell(x,t) = 0$, for all x and t. Then a consumer x will purchase the PS if $1 - (x-t)^2 \geq 0$. Now suppose that the firm targets the entire consumer base. This implies that both $1 - (1-t)^2 - p = 0$, and $1 - t^2 - p = 0$. Therefore $t = \frac{1}{2}$ if $\ell(x,t)$.[13]

Now suppose that the firm faces competition from OSS. Each individual x chooses the PS if

$$1 - (x-t)^2 - (1-x)t^2 \geq \max\{1 - (x-\tau)^2 - (1-x)\tau^2, 0\}.$$

For the lower-critical individual, attention can be restricted to $1 - (x-t)^2 - p - (1-x)t^2 \geq 0$, while attention can be restricted to $1 - (x-t)^2 - p - (1-x)t^2 \geq 1 - (x-\tau)^2 - (1-x)\tau^2$ for the upper-critical individual. Therefore the firm faces demand

$$D(p,t,\tau) = \begin{cases} \dfrac{2(\tau-t)(t+\tau)-p}{(\tau-t)(2+\tau+t)} - \max\left\{0, \dfrac{1}{2}\left(2t + t^2 - \sqrt{4 - 4p - 4t^2 + 4t^3 + t^4}\right)\right\} & \text{if } t < \tau \\ 1 - \dfrac{p + 2(t-\tau)(t+\tau)}{(t-\tau)(2+t+\tau)} & \text{if } t > \tau \end{cases}$$

Fix $t = t^d$ and $\tau = \tau^d$. It follows that the firm sets

[13] Verifying the other boundary conditions shows that profits are indeed maximized by targeting the entire market with $t = \frac{1}{2}$.

$$p(t^d, \tau^d) = \begin{cases} (\tau^d)^2 - (t^d)^2 & \text{if } t^d < \tau^d \\ \frac{1}{2}(t^d - \tau^d)(2 - t^d - \tau^d) & \text{if } t^d > \tau^d. \end{cases}$$

$$t^d(\tau) = \begin{cases} \frac{1}{2}\left(\sqrt{9 + 10\tau + \tau^2} - 3 - \tau\right) & \text{if } t^d < \tau \\ 1 & \text{if } t^d > \tau. \end{cases}$$

To verify that the firm chooses $t^d < \tau^d$, consider the most strict case: $\tau^d = \frac{1}{2}$. Since this is the smallest possible value τ^d takes, if the firm chooses to position itself to the left of the OSS, it will for all other values of $\tau^d > \frac{1}{2}$.

When the firm locates to the left (right) of the OSS, it receives profits of approximately 0.056 (0.0089). Thus the PS is developed with technology

$$t^d\left(\frac{1}{2}\right) = \frac{1}{2}\left(\frac{\sqrt{57}}{2} - \frac{7}{2}\right) < \frac{1}{1 + \sqrt{2}}.$$

Thus far, the analytical example has illustrated the results from propositions 1-7. To illustrate proposition 8, note that when $t^d < \tau^d$,

$$\frac{dt^d(\tau)}{d\tau} = \frac{10 + 2\tau}{4\sqrt{9 + 10\tau + \tau^2}} - \frac{1}{2} > 0, \forall \tau \in \left[\frac{1}{2}, 1\right].$$

As the OSS increases its location, the PS follows suit, as stated in proposition 8.

Extensions

In this section, I consider three broad classes of extensions. The first relates to variations in costs on both the consumer side and the firm side. The second relates to R&D spillovers that occur through the development of OSS and network effects, which can be induced through both installed base size and compatibility between OSS and PS. The third is a market in which OSS is the sole product.

Cost Structures

The learning cost is not essential for the main results. To see this, suppose that $\ell(x, \zeta)$ for all x, ζ, so technology becomes purely horizontal. One of the benefits of contributing to OSS is participating in a project that attracts other high quality programmers. This fact has a very important implication: the location of the OSS will still be above the median due to the fact that contributing to the development of OSS is costly. Therefore the firm still has the incentive to choose a technology level no larger than the median.

The analysis was conducted under assumptions that are least favorable to the results. For example, the firm could choose any location at zero cost, which means it can develop the most technologically advanced software at the same cost as the simplest software. Including this cost would only serve to strengthen the above conclusions. Suppose that the firm faces the strictly increasing and convex cost function $c(t)$, where $c(0) = 0$. No changes occur on the consumers' side of the problem, nor does the firm's pricing decision. The condition outlined in proposition 2 and its associated result becomes

$$\frac{\partial D(p(t^d, \tau^d), t^d, \tau^d)}{\partial t} = c'(t^d).$$

Proposition 6 strengthens since moving slightly to the left of τ^d not only weakly increases the firm's market share, but strictly decreases the firm's costs, leading to a strict increase in profits. Now consider a U-shaped cost function $c(t)$, where $G^{-1}\left(\frac{1}{2}\right) \equiv \operatorname{argmin} c(t)$. Since $\tau^d \geq G^{-1}\left(\frac{1}{2}\right)$, any value $t^d = \tau^d - \kappa$ is strictly preferred to $t^d = \tau^d + \kappa$ for small, positive κ. The argument presented above still holds, as does proposition 6. By invoking the same line of reasoning, the result holds for any U-shaped cost function where $G^{-1}\left(\frac{1}{2}\right) \geq$

argmin $c(t)$. Things are more complicated when $G^{-1}\left(\frac{1}{2}\right) <$ argmin $c(t)$. In this case, the result holds so long as $|c'(t)|$ is sufficiently small at all t. Otherwise, the cost of developing easy to use (low technology) software is excessively costly to the point that it cannot be profitable to do so, regardless of the demand.

Now consider the valuation $v(x, \zeta)$. It may be the case that the sophistication of the software has no impact on the valuation of the software, i.e., a program with technology level ζ gives each individual the same value V, for all x, or $v(x, \zeta) = v(x, x) = \overline{V}$ for all x, ζ (V takes on the maximal value for all individuals), as in [48]. In this case, the results become more extreme. The interpretation of $b(x, \zeta)$ changes slightly, as now there is no direct benefit from use. That is, the contributors have no "need" to be satisfied that is not already satisfied by the PS. Thus the benefits of contributing consist solely of the joy of programming and the delayed benefits (such as signaling).

The first implication of this is in the pricing and location decision of the firm. Since every individual has the same value $\overline{V}$, heterogeneity arises solely in the learning cost. An individual x will purchase the PS over the OSS if

$$\overline{V} - \ell(x, t) - p \geq \max\{\overline{V} - \ell(x, \tau), 0\},$$

which implies that the firm has a dominant strategy in choosing $t = 0$. Thus the assumption that $t = 0$ from [48] is actually an endogenous result of a special case of this model. Since $t = 0$, the learning cost for all individuals is zero. The firm then chooses price to maximize profits, given the OSS location choice τ. Recall that the equilibrium choice of τ is independent of t. Therefore the optimal choice τ^d is unchanged. The firm chooses a price $p(0, \tau^d) > p(t^d, \tau^d)$ and its profits are strictly greater and the catering to ignorance result is more severe, with the firm producing to maximize the value of the most technologically ignorant individual. As a result, the market is fully covered. The location of the OSS remains unchanged while the firm locates closer to the most technologically ignorant individual.

Recall that the OSS developers need not make pre-compiled binaries available. Suppose that there are no binaries available, but the source code is available. It is then up to the individual to download the source code, then compile and install it. This procedure is both time- and knowledge-intensive. Those individuals who are technologically savvy can easily accomplish this task, but those who are not face a higher barrier. Formally, denote $I(x)$ as the compilation/installation cost. It is assumed unchanging in ζ, but strictly decreasing in x, with $I(1) = 0$. An individual x will then choose the PS over the OSS if

$$v(x,t) - p - \ell(x,t) \geq \max\{v(x,\tau) - \ell(x,\tau) - I(x), 0\}. \tag{10}$$

Comparing condition (10) to condition (5), it is obvious that for any given price p, (10) is easier to satisfy. Without changing its price, the firm captures every individual it had when $I(x) = 0$, plus the marginal individuals who, when $I(x) = 0$, consumed the OSS but now purchase the PS. Thus the firm is better off, or in other words, the firm strictly prefers a more costly OSS. This not surprising since a higher cost diminishes the degree of competition the firm faces from the OSS developers by shifting the critical consumer, who is indifferent between choosing PS and OSS, rightward.

Network Effects and R&D Spillovers

It is well known that in industries with network effects, under certain conditions, multiple firms persist in the long run, especially when compatibility is involved [36]. [12] show that when coupled with strategic pricing, compatibility can prevent one product from exhibiting market dominance. It is fairly common in the software world to see compatibility; for example, there is compatibility between Microsoft Office, Google Docs, Apache OpenOffice, and LibreOffice. Similarly, documents written in one of the many LaTeX typesetters are compatible with proprietary editors, such as Scientific Workplace. Thus including Network effects is unlikely to affect the main results. In [10], the demand side learning is related to network externalities, and the authors show that these externalities can lead to only having a single survivor in the long run; however, their model does not include heterogeneous consumers. The remaining papers mentioned in the introduction that discuss network effects do so under the context of firms supporting OSS initiatives. [13] show that the impact of network externalities can lead the either the firm or OSS developers to support network effects, but quality (product differentiation) is modeled as in [48], where the location is treated as exogenous. Thus [13] are unable to illustrate the impact of network externalities on location choice. However, they do show that in the presence of network externalities, both PS and OSS coexist.

R&D spillovers are very important in the context of competition between OSS and PS. These spillovers are often one-sided, moving from the OSS to the PS, and typically depend on the license choice of the OSS. [33] show that when there are large spillovers, both PS and OSS coexist. Counterintuitively, large R&D spillovers occur when the open source license is less open. The GNU GPL license is the most open, and requires that any additions or changes to the source code must also carry the GNU GPL license. This implies that a firm, if using OSS code that is protected by the GNU GPL license, would be required to release their software under the same license, making the software open source. On the other hand, if the OSS is developed under the BSD 2-clause license, then any changes need not be branded under the same license, so the firm can include open source code and the resulting software can still be licensed as PS.

The effect of introducing R&D spillovers to this model is straightforward. Suppose that the OSS is developed under the BSD 2-clause license, so the firm can appropriate any spillovers. Furthermore, assume that the closer the firm locates to the OSS project, the more valuable the spillovers are. The firm would then have the incentive to move its software closer to the OSS, but this incentive is tempered by the demand effect. Moving too close would lead the consumers of the PS to jump ship and choose either the outside option or the OSS. In particular, those far to the left of the PS choose the outside option while those to the right choose the OSS.

Open Source Monopoly and Oligopoly

Very little changes when OSS is the only option. The location decision of the OSS is determined by the system

$$G(\tau^d) - \frac{1}{2} G\left(\hat{x}(\tau^d)\right) = \frac{1}{2}$$

$$b\left(\hat{x}(\tau^d, \tau^d)\right) = K(\hat{x}(\tau^d), \tau^d),$$

which is independent of t. Therefore the OSS's location is independent of competition from PS. What changes is the resulting market share. Instead of the values derived under the mixed duopoly, the critical consumers are defined by the solutions to the equation

$$v(x, \tau) - \ell(x, \tau) = 0$$

when interior, or the associated boundary conditions when either $v(1, \tau) - \ell(1, \tau) > 0$ or $v(0, \tau) - \ell(0, \tau) > 0$. Contributors' motivations include solving one's own needs, intellectual curiosity, and the joy of programming. These intrinsic motivations support the notion that an individual’s decision on whether or not to contribute, and where the individual's preferred location is, is independent of the PS’s location. Maximizing market share/profits is a rarely cited reason and is typically reserved for commercial developers of OSS.

Independence, as defined above, need not always be the case. In many instances, contributions to and the direction of OSS are motivated (often financially) by for profit entities. Take the Linux-based Fedora operating system. Red Hat Inc. has veto power over all of Fedora's development decisions. While clearly not a monopoly, this example illustrates the important points, which still hold under a monopoly. This idea can be included directly into the associated framework by modeling Red Hat as an atomic agent in the bargaining problem defined in analysis. While each individual developer has zero bargaining power, Red Hat has enough bargaining power to influence the outcome. Since releases of Red Hat Enterprise Linux are branches of the Fedora operating system, Red Hat has a strong incentive to exert this influence. However, in OSS guided solely by community interests, the development decisions remain independent of competition from PS.

On the other hand, introducing competition from other OSS will change the locations. The community is divided and each project locates at the median point of its developers' interests. This closely resembles the development of LaTeX typesetters. There are many open source options, such as TexMaker, TexStudio, Emacs, each varying in its capabilities and ease of use.

Discussion

Product differentiation arises endogenously through competitive forces, where the developers of the PS cater to the technologically ignorant consumers while the developers of OSS cater to the technologically elite consumers. Catering to ignorance should not be seen as a surprise. It may be less expensive to develop simple software as opposed to more complex software with many technical features (though it need not be). Furthermore, there are more individuals who are technologically ignorant than technologically elite. In the modeling setup, I assume an equal measure of technologically ignorant individuals and technologically elite individuals. This assumption represents the strictest setting; if the firm has the incentive to target the less sophisticated individuals, then that incentive only increases as the measure of less sophisticated individuals increases.

The mere presence of OSS pushes the PS further away from the technologically elite consumers, towards the technologically ignorant consumers. These results have significant competitive and policy implications. For instance, if entry by an OSS competitor cannot be blocked, a firm developing PS may have the incentive to "push" the OSS towards the most elite individuals, e.g., through a steering committee as in the Mozilla example. Doing so would allow the market to closely resemble the monopoly market.

[45] shows that governments around the world have been implementing policies in support of OSS, but claims that it is unclear whether this is efficiency improving. Efficiency should be increased through the entry of OSS; however, it is unclear whether or not costly government investment would be beneficial. Government investment could either help solve the potential free-rider problem that exists when there is community-based development, or it could crowd out investments made by firms to OSS entities, like those discussed in [37] and [46]. [19] provide evidence indicating that the supply side of OSS

benefits from the enforcement of intellectual property rights, while some regulation has led to an increase on the number of OSS developers.

Some OSS projects are developed because those technologically elite individuals felt that the PS was designed not for them, but for the technologically ignorant consumers. In response to the firm's behavior, OSS is developed for the elite. However, this development further pushes the PS towards the technologically ignorant consumers. This process becomes a self-propelling mechanism, where the firm's location decision induces competition from OSS, which in turn induces the firm to reevaluate its location, which induces more OSS competition, *ad infinitum*. The previous illustration is an extrapolation of the static model as a dynamic process, where entry occurs over time and what follows is a firm located near $t = 0$, with OSS located at various intervals across the spectrum of technological savviness. The degree to which the firm targets the most ignorant consumer is increasing in how costly the OSS is, relative to the PS. If the consumer must compile the software herself, then the cost of using the OSS increases. It is, as if, the OSS moves further to the right. The firm is thus able to capture the critical individuals who would have used the OSS had they not faced this added cost.

Conclusion

There is still much work to be done in studying competition in the software industry. OSS is not the same as free software. Many companies offer free OSS and sell complementary services, such as customer and user guides. In some cases, such as the Red Hat Enterprise Linux distribution, firms sell the actual software at a price. While any individual can obtain any OSS for free (monetarily), there may be a significant opportunity cost in doing so, typically in terms of time and effort. Thus one could model OSS by considering a firm practicing second-degree price discrimination, where consumers with a low opportunity cost choose to download the source code and expend costly effort to convert it into usable software, while those with a higher opportunity cost of time pay a premium to receive the same OSS in a more readily usable condition. This price discrimination leaves the OSS developers with an interesting strategic decision. Is there an optimal ``difficulty'' the developers should set that maximizes the developers' profits, and how does this compare to the level that maximizes total welfare?

References

1. Ajila, S.A. and Wu, D.Empirical Study of the Effects of Open Source Adoption on Software Development Economics. *Journal of Systems and Software*, *80*, 9 (September 2007), 1517–1529.

2. Aksulu, A. and Wade, M.A Comprehensive Review and Synthesis of Open Source Research. *Journal of the Association for Information Systems*, *11*, 11/12 (November 2010), 576–656.

3. Anderson, S., de Palma, A., and Thisse, J.-F.Privatization and efficiency in a differentiated industry. *European Economic Review*, *41*, 9 (1997), 1635–1654.

4. Athey, S. and Ellison, G.Dynamics of Open Source Movements. *Journal of Economics and Management Strategy*, *23*, 2 (2014), 294–316.

5. Birke, D.The Economics of Networks: A Survey of the Empirical Literature. *Journal of Economic Surveys*, *23*, 4 (2009), 762–793.

6. Bitzer, J., Schrettl, W., and Schröder, P.J.H.Intrinsic Motivation in Open Source Software Development. *Journal of Comparative Economics*, *35*, 1 (March 2007), 160–169.

7. Bitzer, J.Commercial versus open source software: the role of product heterogeneity in competition. *Economic Systems*, *28*, 4 (2004), 369–381.

8. Bulow, J.Durable-Goods Monopolists. *Journal of Political Economy*, *90*, 2 (April 1982), 314–332.

9. Bulow, J.An Economic Theory of Planned Obsolescence. *Quarterly Journal of Economics*, *101*, 4 (November 1986), 729–750.

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

11. Casadesus-Masanell, R. and Llanes, G.Mixed Source. *Management Science*, *57*, 7 (2011), 1212–1230.

12. Chen, J., Doraszelski, U., and Harrington Jr., J.E.Avoiding Market Dominance: Product Compatibility in Markets with Network Effects. *RAND Journal of Economics*, *40*, 3 (2009), 455–485.

13. Cheng, H., Liu, Y., and Tang, Q.The Impact of Network Externalities on the Competition Between Open Source and Proprietary Software. *Journal of Management Information Systems*, *27*, 4 (2011), 201–230.

14. Cho, S. and Duggan, J.Bargaining foundations of the median voter theorem. *Journal of Economic Theory*, *144*, 2 (2009), 851–868.

15. Christiaans, T.On the Dynamics of Competition between Commercial and Free Software. *Studies in Microeconomics*, *1*, 1 (2013), 37–58.

16. Dey, D., Lahiri, A., and Liu, D.Consumer Learning and Time-Locked Trials of Software Products. *Journal of Management Information Systems*, *30*, 2 (2013), 239–268.

17. Economides, N. and Katsamakas, E.The Economics of Open Source Software Development. In J. Bitzer and P.J.H. Schröder, eds., Elsevier B.V., 2006, pp. 207–218.

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

19. Engelhardt, S. and Freytag, A.Institutions, culture, and open source. *Journal of Economic Behavior and Organization*, *95*, (2013), 90–110.

20. Fang, Y. and Neufeld, D.Understanding Sustained Participation in Open Source Software Projects. *Journal of Management Information Systems*, *25*, 4 (2009), 9–50.

21. Fershtman, C. and Gandal, N.Open Source Software: Motivation and Restrictive Licensing. *Journal of International Economics and Economic Policy*, *4*, 2 (August 2007), 209–225.

22. Fershtman, C. and Gandal, N.A Brief Survey of the Economics of Open Source Software. *CEPR Discussion Paper No. DP8434*, (June 2011).

23. De Fraja, G. and Delbono, F.Game Theoretic Models of Mixed Oligopoly. *Journal of Economic Surveys*, *4*, 1 (1990), 1–17.

24. Hars, A. and Ou, S.Working for Free? - Motivations of Participating in Open Source Projects. In *Proceedings of the 34th Hawaii International Conference on System Sciences*. 2001.

25. Holmström, B.Managerial Incentive Problems: A Dynamic Perspective. *Review of Economic Studies*, *66*, 1 (1999), 169–182.

26. Jaisingh, J., See-To, E.W.K., and Tam, K.Y.The Impact of Open Source Software on the Strategic Choices of Firms Developing Proprietary Software. *Journal of Management Information Systems*, *25*, 3, 241–275.

27. Johnson, J.P.Open Source Software: Private Provision of a Public Good. *Journal of Economics and Management Strategy*, *11*, 4 (2002), 637–662.

28. Von Krogh, G. and von Hippel, E.The Promise of Research on Open Source Software. *Management Science*, *52*, 7 (July 2006), 975–983.

29. Kumar, V., Gordon, B.R., and Srinivasan, K.Competitive Strategy for Open Source Software. *Marketing Science*, *30*, 6 (2011), 1066–1078.

30. Lerner, J., Pathak, P.A., and Tirole, J.The Dynamics of Open-Source Contributors. *American Economic Review*, *96*, 2 (May 2006), 114–118.

31. Lerner, J. and Tirole, J.The Open Source Movement: Key Research Questions. *European Economic Review*, *45*, 4-6 (May 2001), 819–826.

32. Lerner, J. and Tirole, J.Some Simple Economics of Open Source. *Journal of Industrial Economics*, *L*, 2 (June 2002), 197–234.

33. Llanes, G. and de Elejalde, R.Industry Equilibrium with open-source and proprietary firms. *International Journal of Industrial Organization*, *31*, 1 (2013), 36–49.

34. Matsumura, T.Partial privatization in mixed duopoly. *Journal of Public Economics*, *70*, 3 (1998), 473–483.

35. Maurer, S.M. and Scotchmer, S.Open Source Software: The New Intellectual Proprety Paradigm. *NBER Working Paper Series*, (2006).

36. Mitchell, M.F. and Skrzypacz, A.Network Externalities and Long-Run Market Shares. *Journal of Economic Theory*, *29*, 3 (November 2006), 621–648.

37. Mustonen, M.When Does a Firm Support Substitute Open Source Programming? *Journal of Economics and Management Strategy*, *14*, 1 (2005), 121–139.

38. Myatt, D.P. and Wallace, C.An Evolutionary Analysis of the Volunteer's Dilemma. *Games and Economic Behavior*, *62*, 1 (January 2008), 67–76.

39. Nadeau, T.Learning from Linux. 1999. http://www.os2hq.com/archives/linmemo1.htm.

40. Noyes, K.Open Source is Taking Over the Software World, Survey Says. 2013. http://www.pcworld.com/article/2035651/open-source-is-taking-over-the-software-world-survey-says.html.

41. Open Source Initiative.GNU General Purpose License v3. http://opensource.org/licenses/GPL-3.0.

42. Open Source Initiative.The BSD 2-Clause License. http://opensource.org/licenses/BSD-2-Clause.

43. Open Source Initiative.The Open Source Definition. http://opensource.org/osd.

44. Pirillo, C.The Future Of Software is Open Source. 2008. http://chris.pirillo.com/the-future-of-software-is-open-source/.

45. Raju, K.D.Is the Future of Software Development in Open Source Software? Proprietary vs. Open Source Software: A Cross Country Analysis. *Journal of Intellectual Property Rights*, *12*, 2 (March 2007), 1–20.

46. Riehle, D.The Commercial Open Source Business Model. *Lecture Notes in Business Information Processing*, *36*, (2009), 18–30.

47. Saurer, R.M.Why Develop Open-Source Software? The Role of Non-Pecuniary Benefits, Monetary Rewards and Open-Source License Type. *Oxford Review of Economic Policy*, *23*, 4 (2007), 605–619.

48. Sen, R.A Strategic Analysis of Competition between Open Source and Proprietary Software. *Journal of Management Information Systems*, *24*, 1 (2007), 233–257.

49. Shy, O.*The Economics of Network Industries.* Cambridge: Cambridge University Press, 2001.

50. Shy, O.A Short Survey of Network Economics. *Review of Industrial Organization*, *38*, 2 (March 2011), 119–149.

51. Spence, M.Monopoly, Quality, and Regulation. *Bell Journal of Economics*, *6*, 2 (1975), 417–429.

52. Waldman, M.Durable Goods Pricing When Quality Matters. *Journal of Business*, *69*, 4 (October 1996), 489–510.

Appendix: Proofs

Lemma 1: Proof. To the contrary, suppose not. This implies that there exists three individuals x_1, x_2, and x_3, with $0 \leq x_1 < x_2 < x_3 \leq 1$ such that

$$v(x_1, t^m) - p - \ell(x_1, t^m) > 0$$
$$v(x_2, t^m) - p - \ell(x_2, t^m) < 0$$
$$v(x_3, t^m) - p - \ell(x_3, t^m) > 0$$

By the intermediate value theorem, there must exist some $w \in (x_1, x_2)$ such that $v(w, t^m) - p - \ell(w, t^m) = 0$. Similarly, there must be some $z \in (x_2, x_3)$ such that $v(z, t^m) - p - \ell(z, t^m)$. This implies that there exists some point $y \in (w, z)$ such that $v_x(y, t^m) - \ell_x(y, t^m) = 0$ and $v_{xx}(y, t^m) - \ell_{xx}(y, t^m) > 0$, a contradiction since $v(x,t) - \ell(x, t^m)$ is strictly concave in x. Thus the support of the demand is a non-degenerate convex subset of [0,1]. ***Q.E.D.***

Lemma 2: Proof. Recall that $v(x,t) - \ell(x,t) - p$ is strictly concave and single peaked. Set $t = t^m$. For prices p and p', $p \neq p'$, the difference between $v(x, t^m) - \ell(x, t^m) - p$ and $v(x, t^m) - \ell(x, t^m) - p'$ is constant and equal to $p' - p$. At $p = 0$, $v(x, t^m) - \ell(x, t^m). 0$ for a subset of individuals. Thus for a fixed t, it must be that the indifferent individual on each side of t must (weakly) move closer to t as p increases. ***Q.E.D.***

Proposition 1: Proof. The objective function can be written as $\max_p p \int_{\underline{x}(p,t)}^{\overline{x}(p,t)} g(x)dx$. The first-order necessary condition is

$$\int_{\underline{x}(p^*,t)}^{\overline{x}(p^*,t)} g(x)dx + p^* \left[g\big(\overline{x}(p^*,t)\big) \frac{\partial \overline{x}(p^*,t)}{\partial p} - g\left(\underline{x}(p^*,t)\right) \frac{\partial \underline{x}(p^*,t)}{\partial p} \right] = 0$$

The second-order sufficient condition is satisfied by the log-concavity assumption. By the implicit function theorem, $p^* = p(t)$. The first-order necessary condition can thus be rewritten as

$$\frac{1}{D(p(t),t)} \left(1 + p(t) \frac{\frac{\partial D(p(t),t)}{\partial p}}{D(p(t),t)} \right) = 0.$$

Note that $D(p(t),t)$ must be positive in equilibrium for any chosen t. The result follows directly. ***Q.E.D.***

Proposition 2: Proof. The objective function can be written as $\max_{t \in [0,1]} p(t) D(p(t),t)$. Note that by the theorem of the maximum and concavity of the utility function, $p(t)$ is continuously differentiable about $t = t^m$ if $t^m \in (0,1)$. If $t^m \in \{0,1\}$, then I define

$$p'(t) \equiv \begin{cases} \lim\limits_{\Delta \to {}_{+}0} \dfrac{p(\Delta) - p(0)}{\Delta} & \text{if } t = 0 \\ \lim\limits_{\Delta \to {}_{-}1} \dfrac{p(1) - p(1-\Delta)}{1 - \Delta} & \text{if } t = 1, \end{cases}$$

The first-order necessary condition for an interior solution can be written as

$$t^m \frac{\frac{\partial D(p(t^m), t^m)}{\partial t}}{D(p(t^m), t^m)} = \mu(t^m) = 0.$$

To rule out $t^m = 0$, it must be that $\frac{\partial D(p(t^m),t^m)}{\partial t} \leq 0$, for $t^m = 0$, which cannot be true. To see this, note that at $t^m = 0$, $\underline{x}(p(0), 0) = 0$ since this product is most valuable to the consumer located at $x = 0$. For a small enough $\varepsilon > 0$, $\underline{x}(p(0), \varepsilon) = 0$. However, any individual to the right of $t^m = \varepsilon$ is strictly better off than when $t^m = 0$. Thus $\overline{x}(p(0), \varepsilon) > \overline{x}(p(0), 0)$, which means $\frac{\partial D(p(0),0)}{\partial t} \equiv \lim_{\Delta \to +0} \frac{D(p(0),\Delta) - D(p(0),0)}{\Delta} > 0$. An identical argument holds for $t^m = 1$. ***Q.E.D.***

Proposition 3: Proof.
Lemma 4 *Suppose* $\ell(x, \zeta) = 0$ *for all* x, ζ. *Then* $t^m = G^{-1}\left(\frac{1}{2}\right)$.
Proof. From Proposition 2, t^m exists and is both unique and interior. Suppose $t^m > G^{-1}\left(\frac{1}{2}\right)$. Then, by log concavity of $G(\cdot)$, there exists an $\epsilon > 0$ such that

$$\int_{\underline{x}(p(t^m),t^m-\epsilon)}^{\overline{x}(p(t^m),t^m-\epsilon)} g(y)dy \geq \int_{\underline{x}(p(t^m),t^m)}^{\overline{x}(p(t^m),t^m)} g(y)dy$$

The inequality is strict for all non-uniform log-concave distributions. Thus $t^m > G^{-1}\left(\frac{1}{2}\right)$ cannot possibly be an equilibrium. A symmetric argument holds for $t^m < G^{-1}\left(\frac{1}{2}\right)$ by replacing each $-\epsilon$ with $+\epsilon$. ***Q.E.D.***

To see this result, recall that $\ell(x, t)$ is strictly increasing and convex in t. Thus making an arbitrarily small increase in t leads to an increase in the learning cost and this increase is itself increasing in t. Therefore the reverse movement (a slight decrease in t) leads to a decrease in costs, with that decrease decreasing in magnitude the larger the change. However, the effect on value is symmetric. Take two individuals x_1 and x_2 and the median technology level t, where $x_1 < t < x_2$ that satisfies $v(x_1, t) = v(x_2, t)$. It follows for a small, positive value v that $v(x_2, t + v) - v(x_2, v) = v(x_1, t - v) - v(x_1, t)$. Thus a decrease in t has a stronger demand effect than an increase in t by the same amount, *ceteris paribus*. This implies the existence of a value $t^* = t - v$ such that, for the price p that satisfies $\varepsilon\big(p(t)\big) = -1$, $p(t)D(p(t), t) < p(t)D(p(t), t^*)$. Therefore the equilibrium must be less than the median. ***Q.E.D.***

Proposition 4: Proof. Recall that firm demand is positive only if

$$v(x, t^d) - p(t^d, \tau^d) - \ell(x, t^d) \geq v(x, \tau^d) - \ell(x, \tau^d),$$

for a positive mass of individuals. At $t^d = \tau^d$ the above simplifies to $-p(t^d, \tau^d) \geq 0$, which is false for all individuals. ***Q.E.D.***

Proposition 5: Proof. The proof is identical to that of Proposition 1, making the appropriate substitution for demand. ***Q.E.D.***

Lemma 3: Proof. By assumption, since $b(1, \tau) > K(1, \tau)$, the most technologically sophisticated individual is willing to contribute. Since $b(x, \tau)$ is weakly concave in x and $K(x, \tau)$ is strictly convex in x, there must be a positive mass of individuals with $b(x, \tau) \geq K(x, \tau)$. Thus there exists a value $\hat{x}(\tau)$ such that $b(\hat{x}(\tau), \tau) = K(\hat{x}(\tau), \tau)$. ***Q.E.D.***

Proposition 6: Proof. Consider the most extreme case: $\tau^d = G^{-1}\left(\frac{1}{2}\right)$. If the firm locates at $t^d > G^{-1}\left(\frac{1}{2}\right)$, then the the firm's quantity demanded at $p = 0$ is strictly less than $\frac{1}{2}$, while the firm's quantity demanded at $t^d < G^{-1}\left(\frac{1}{2}\right)$ and $p = 0$ is strictly greater than $\frac{1}{2}$. Also, note that utility is linear in price, which implies that the indifference condition is both continuous and linear in price. Thus for any price p, there exists a $t^d < G^{-1}\left(\frac{1}{2}\right)$ such that for any $t > G^{-1}\left(\frac{1}{2}\right)$, $pG\left(\underline{x}(p, t^d, \tau^d)\right) \geq p\left[1 - G\left(\overline{x}(p, t^d, \tau^d)\right)\right]$. Therefore the firm strictly prefers to locate to the left of the open source software. ***Q.E.D.***

Proposition 7: Proof. Denote π^m as the firm's profits under the monopoly and π^d as the firm's profits under the mixed duopoly. To the contrary, suppose that $t^m < t^d < \tau^d$. Under the mixed duopoly, the firm obtains zero demand from consumers located in the neighborhood about $x = \tau^d$ and zero demand from all values of $x \geq \tau^d$. If $\pi^m > \pi^d$, then the firm in the mixed duopoly can replicate the monopolist and see a strict increase in profits. Thus $t^d > t^m$ is not an equilibrium when $\pi^m > \pi^d$. Similarly, if $\pi^d > \pi^m$, then the monopolist can replicate the duopolist and thus $t^m < t^d$ is not an equilibrium. Furthermore, it cannot be that $t^m < t^d$ and $\pi^m = \pi^d$ since the monopolist could set $t^m = t^d$, and see a strict increase in profits since it faces no competition from an open source option. ***Q.E.D.***

Proposition 8: Proof. The result immediately follows from Proposition 7.

Figure 1. Illustration of the cost of learning:
(a) as a function of the technological sophistication ζ.
(b) as a function of the individual prowess x.

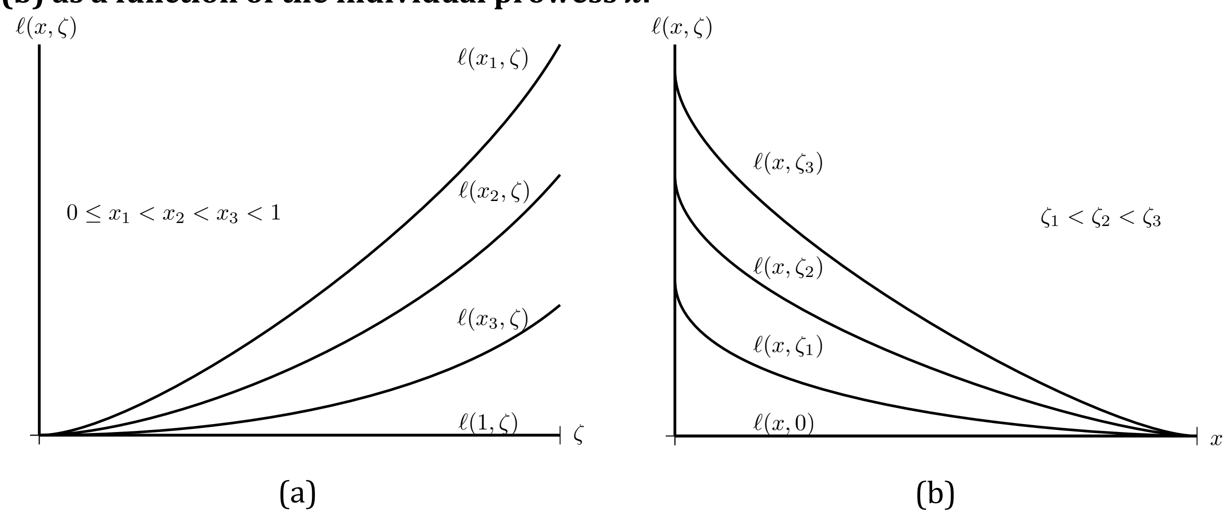

Figure 2. Illustrating the subgame perfect equilibrium under the monopoly

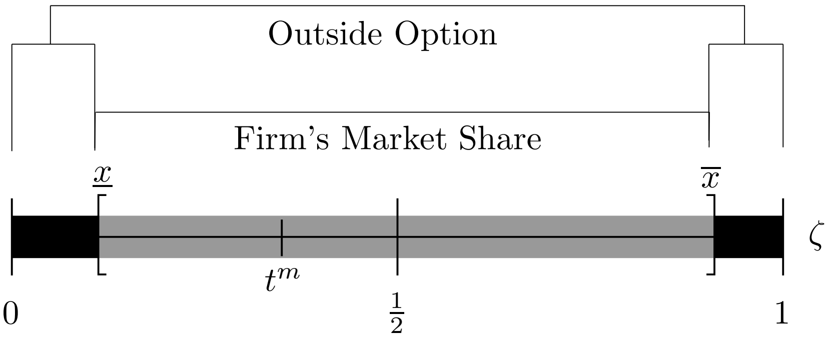

Figure 3. Location of the individual indifferent between contributing to the OSS and not contributing.

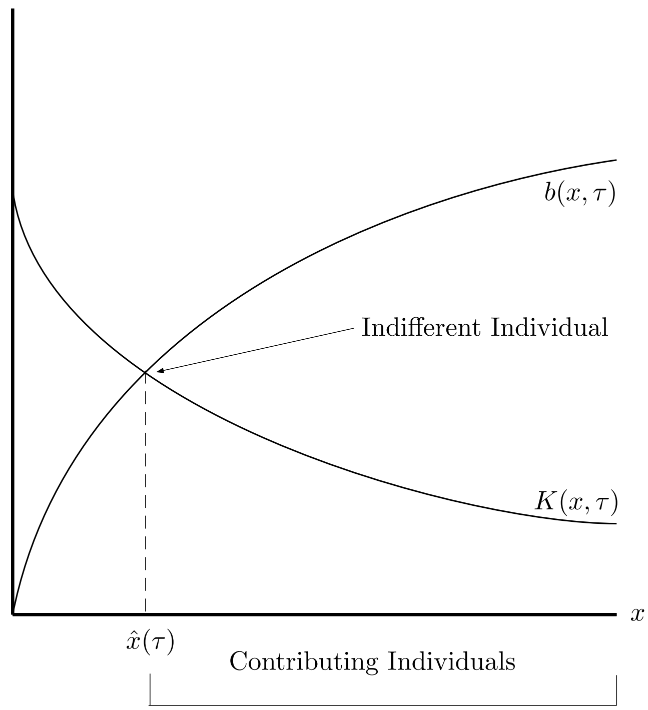

Figure 4. Illustration of the subgame perfect equilibrium under a duopoly