Piracy in Commercial vs. Open-Source Software Competition*

Arne Rogde Gramstad†
University of Oslo
May 15, 2014

Abstract

I develop a model to study competition between a commercial software provider and a free-of-charge open-source software substitute in the presence of network externalities and software piracy. Whenever the commercial software product is endowed with a large installed base of users relative to the open source competitor, it may be feasible to strategically set the price such that a critical mass of users is not attainable for the open source product. I coin this pricing strategy "network strategic pricing". The commercial software provider will take the entire relevant market whenever network strategic pricing is imposed. When piracy is included in the model, network strategic pricing turns out to be more likely to be the optimal pricing strategy relative to the no-piracy benchmark. This is because the cost of imposing network strategic pricing (i.e., the price effect) is lower when piracy is present. I also provide empirical evidence suggesting that software piracy has a negative impact on usage of the open-source operating system Linux.

JEL classification: L13; L17; O34

Keywords: Piracy, open-source software, pricing strategies, duopoly

*I would like to thank Torben Mideksa, Tore Nilssen and Espen Stokkereit for helpful comments and suggestions.
†Department of Economics, University of Oslo, a.r.gramstad@econ.uio.no
1 Introduction

This paper analyzes optimal pricing strategies for a seller of a commercial software product facing competition from a free-of-charge open-source software product and illegal copies of the commercial product. The primary objective is to understand how incentives for strategically using network externalities in the pricing decision are affected by the presence of piracy.

According to the 2011 BSA Global Piracy study an estimated 42% of all commercial software used in 2011 was illegally pirated (Business Software Alliance, 2012). In addition to the substantial impact software piracy is likely to have on the profitability of software firms, the market structure will necessarily be affected given the size of this illicit economy.

The entry of successful free and open-source software (FOSS) products the recent decades has transformed the information technology industry. As the world has become increasingly connected, collaboration in the development of open-source software has been immensely facilitated. A wide range of free-of-charge software products are now able to compete with their proprietary counterparts, and have even come to dominate certain segments of the software markets, such as web servers (Apache) and operating systems for mobile phones (Android).

Commercial software products still dominate other segments of software markets, despite the availability of free-of-charge open-source substitutes. Office suites like OpenOffice and LibreOffice are far less popular than the dominant Microsoft Office in user shares, and less than 2% of desktop computers run a Linux-based operating system.¹

These large differences in market shares are likely to be explained by network externalities. That is, the value to consumers of some software products increases in the size of the installed base. Products may exhibit network externalities, either directly if individuals prefer to use the same products as those they interact with, or indirectly if the quantity and quality of compatible applications is increasing in the installed base.²

²See e.g. Katz and Shapiro (1985) for a general treatment of network externalities. Products where the quantity of compatible application increases in the installed base are typically "two-sided" in the sense that there exists positive externalities between consumers and developers of compatible applications, see e.g. Rochet and Tirole (2006).
consumers demand the size of the installed base to exceed a threshold in order to be willing to obtain the good. In that case, there exists a so-called critical mass of users necessary for self-sustained adoption a product.  

Assuming the existence of a critical mass, this paper identifies a pricing strategy in which the seller of the commercial product deliberately sets the price such that a critical mass of users is not attained for the open source competitor, a strategy coined "network strategic pricing". In turn, the entire market is captured by the commercial software product. The optimality of imposing such a strategy depends among other things on the initial size of the installed bases of the products and the price reduction relative to "passive" pricing where both products have installed bases exceeding the critical mass.

When pirated versions of the commercial product are available to consumers, "network strategic pricing" is more likely to be imposed. The explanation is two-fold: First, piracy increases the size of the installed base of commercial software. Second, competition from illegal copies causes downward price pressure of legal copies. These two effects reduce the price reduction from "passive" pricing to "network strategic pricing" where the critical mass of the installed base is not attained for the open source product.

Therefore, based on the model presented in this paper, it can be argued that piracy may significantly affect the way commercial software vendors react to competition from free and open-source software products, and that the price reduction following competition from software piracy is too large to be explained by piracy alone.

The observed behavior of seemingly low prices offered major software firms may partly be explained by the model in the present paper. In the United States v. Microsoft Corp. (2001) antitrust case, expert witnesses argued that the pricing of Microsoft Windows was "too low (...) to be legally classified as a monopoly." Although Microsoft exercises price discrimination of their products, low prices, especially to the non-business segment, remain today. For instance, the Windows 8 license to original equipment manufacturers (OEMs) is currently priced at just $15. One explanation for this pricing

---

3 Economides and Himmelberg (1995) define a critical mass as the smallest network size supported in equilibrium.


behavior may be the threat of successful adoption of competing open-source products.

Open-source software is the generic term for a class of software products released under a public copyright license, i.e., the users (or licensees) are given the right to freely copy or change the product. The source code (the "recipe") is publicly available, and the product will therefore usually be free of charge once a software product is released. Open source licenses may be "restrictive" in the sense that licensees are restricted to release copies or modified versions under the same license as the original version, which ensures that the product remains open source (e.g. the General Public License (GPL)). Open source licenses can also be "unrestrictive" in the sense that licensees are more or less unrestricted in the type of license modified and copied versions can take (e.g. the Berkeley Software Distribution (BSD) license).

Development of open source software happens both through contributions by volunteers and in firms. Lerner and Tirole (2002) argue that voluntary contributions to open source development may be explained by peer recognition and job market signaling, while a survey by Lakhani and Wolf (2005) finds that enjoyment-based motivation is the most important factor. An extensive survey by Lerner and Schankerman (2010) suggests that there exist economies of scope between open source and closed source software development, explaining why open source development is more likely to take place in large firms. Also, firms engaging in development of customized software are more likely to develop under an open source license. In general, there might be cost savings and qualitative benefits for firms developing open source software since users can help identify bugs in the code as well as contributing to the development themselves. However, by developing open source software, for-profit firms must be able to generate revenue from complementary products such as software applications, advertising, and customer support. Firms may also benefit from releasing unsuccessful proprietary (closed source) software products under an open source license.\textsuperscript{6}

\textit{Related literature:} The present paper relates to previous research on pricing of goods exhibiting network externalities as well as the literature concerning software piracy and commercial vs. open-source competition specifically. The concept of strategically taking advantage of network externalities in the pricing decision is explored by among others Fudenberg and Tirole (2000). They assume that the market has room for only one

\textsuperscript{6}The web browser Mozilla Firefox and the office suite OpenOffice are both examples of initial proprietary products re-released under open source licenses (from Netscape Navigator and StarOffice, respectively). See Caulkins et al. (2013) for a theoretical approach to such firm behavior.
vendor of a network good, and that the incumbent may find it optimal to set a low price in order to discourage entry of a firm selling an incompatible network good. Cabral (2011) also finds that firms with large networks may choose to reduce the price. This is because larger firms have more to gain from increasing the network size, and may thus cut the price if the concern for future profits dominates present profits.

This paper is most closely related to the literature on pricing in commercial vs. open source software competition. The contribution by Casadesus-Masanell and Ghemawat (2006) is most likely the paper closest to the present one as they not only address commercial vs. open source software competition, but also extend the analysis to include software piracy. In a dynamic Windows vs. Linux duopoly model with network externalities they show that Windows may use a forward-looking pricing strategy ensuring that the market does not tip to Linux dominance, despite an assumed competitive advantage to Linux from user-development. By extending the model to an exogenously given piracy rate of Windows they find larger steady-state differences in user shares between Windows and Linux. However, without an endogenous piracy rate, the scope of their analysis is limited.

Economides and Katsamakas (2006) consider a model of two-sided platform competition between a proprietary and an open source platform, applicable to e.g. operating systems with compatible proprietary applications. By assuming higher relative switching costs to the open source platform as well as low substitutability between platforms, they find that the proprietary platform is more likely to dominate the market.

While not explicitly modeling two-sided platforms Bitzer (2004) reaches a similar conclusion, that survival of proprietary software products is conditional on maintaining a high degree of product heterogeneity. In my own model I find that a high degree of heterogeneity (or product differentiation) is likely to yield an equilibrium with coexistence of the two products without piracy, while high product differentiation combined with the existence of piracy may incentivize the developer of the proprietary product to impose a pricing strategy that causes the proprietary software product to dominate the market.

Other papers explicitly take investment and development into consideration. Lanzi (2009) considers a duopoly model where the proprietary provider maximizes profits, while the open source community maximizes the number of users. In addition, he assumes that the technological skill level of consumers differ, and that higher skills are
needed to be able to use open source software. He finds that the market outcome depends on the difference in skill-level needed to be able to operate open source software compared to the proprietary competitor.

Lee and Mendelson (2008) model open source development as a function of consumer surplus and intrinsic value to contributors. The competitive outcome and level of investment depend among other things on timing and product compatibility. One result is that the potential threat from open source may induce the commercial software firm to increase investments in development and lower prices.

Mustonen (2003) provides a model where programmers choose whether to be employed by a commercial software firm or join the open source community to signal their skills. Co-existence of the two types of software products depends on implementation costs and whether the commercial software firms can attract a sufficiently large share of the programmers. He finds that market outcomes with dominance by both open source and commercial software may occur.

In the theoretical literature on software piracy it is argued that software piracy can be used as a commitment to intertemporal price discrimination (Takeyama, 1997), as a device for signaling of quality (Takeyama, 2003), and as a method of by-passing antitrust regulation on predatory pricing (Ben-Shahar and Jacob, 2004). Moreover, the impact from piracy on price and quality may differ substantially between short run and long run equilibria (Bae and Choi, 2006). Survey literature suggests that among students, price is the most important motivation for obtaining pirated products over legal copies (Cheng et al., 1997), while norms and fear of punishment are of equal importance among working adults (Peace et al., 2003).

Piracy and competition is addressed by among others Shy and Thisse (1999), Jain (2008), and Rasch and Wenzel (2013). This work suggests that limited copy protection can be optimal for the software industry if network externalities are strong (Shy and Thisse, 1999), and may be beneficial to firms even in the absence of network externalities since piracy might reduce price competition (Jain, 2008). Rasch and Wenzel (2013) address the trade-off between user surplus and software developers’ surplus of copy protection in two-sided platform competition. For a general treatment on piracy of digital products, excellent literature reviews are provided by Peitz and Waelbroeck (2006) and Belleflamme and Peitz (2012).

The rest of the paper is structured as follows: Part 2 provides the benchmark model
without software piracy, while the model is extended to include piracy in part 3. Part 4
discusses the model’s implications by comparing the results from the models in the two

2 Benchmark model: No piracy

As a benchmark, competition is analyzed without the availability of illegal copies. The
first benchmark model is a simple location model in which a commercial software provider
faces competition from a differentiated free-of-charge substitute. Network externalities
are added to the model by imposing a restriction of users requiring that the total user
share of a given software product must exceed a given critical mass in order to obtain
the product.

There are two types of software products available, $W$ (as in Windows) and $L$ (as
in Linux). As the choice of notation suggests, $W$ is the commercial type and is sold at
some price $p$, while $L$ is the open source type which is free of charge. The two products
are substitutes, and customers can only choose to obtain one of the two types. Hence,
there exist two products, but only one price.

We consider a population of software users of size normalized to one. There are two
types of software users: new users, who make up a share $\sigma$ of the population, and old
users with a share of $1 - \sigma$. The new users will have to choose between purchasing $W$
and downloading $L$ for free. The old users are assumed to have made their decisions
on which product to obtain earlier, and will not revise their decision. Hence, it is not
possible to extract further profits from the old user population.

Among the old user population a share $(1 + \lambda)/2$ belong to the $W$ network, while the
rest $(1 - \lambda)/2$ belong to the $L$ network. Since user shares only take non-negative values,
we have $\lambda \in [-1, 1]$.

The measures of old users of $W$ and $L$ are therefore given by

$$W^{old} = (1 - \sigma)(1 + \lambda)/2$$

and

$$L^{old} = (1 - \sigma)(1 - \lambda)/2,$$

respectively. Clearly, if $\lambda = 1$ the entire old user population will be $W$ users. Likewise,
$\lambda = -1$ means that all of the old users are $L$ users.
$W^{\text{old}}$ and $L^{\text{old}}$ can be interpreted as an exogenous installed base of each software product. For instance, as of January 2014 approximately 30% of all desktop computers run Windows XP, an operating system released in 2001. These numbers suggest there is a population of significant size contributing to the network size of all Windows versions, while having a negligible direct effect on Microsoft’s revenue. Whenever network externalities are assumed away, the old users will not affect prices and can therefore be disregarded.

### 2.1 No network externalities

New users, the section of the population who choose between purchasing $W$ or downloading $L$ for free, are uniformly distributed over a line of length $t$. $W$ is located at $-t/2$, and $L$ is located at $t/2$, where $t > 0$. The valuation to the median consumer is denoted $v_W$ and $v_L$ for each software type. A taste variable $\omega$ uniformly distributed over $[-t/2, t/2]$, represents heterogeneity in preferences among consumers. Payoffs for a consumer with taste $\omega$ from purchasing $W$ and downloading $L$ are given by $U_W = v_W - p - \frac{1}{2}\omega$ and $U_L = v_L + \frac{1}{2}\omega$, respectively.

An increase in $t$ will on average increase (decrease) the benefit of using $W$ ($L$) for any consumer with a taste variable $\omega < 0$. To take a concrete example this parameter can measure the differentiation between the product with regards to simplicity and customizability. Consumers with preferences leaning towards closed source software might prefer a simple system with little room for customizing, while consumers leaning towards open-source software will increase their benefit when the system becomes more customizable.\(^8\) Note that there is only one price given despite the existence of two products. A new user will choose to purchase $W$ at price $p$ over downloading $L$ for free if $U_W \geq U_L$, i.e., whenever:

$$(v_W - p) - v_L \geq \omega,$$

which implies that $L$ is preferred to $W$ when the opposite is true:

$$(v_W - p) - v_L < \omega.$$
Figure 1: User shares among new users for a given level of quality and price

The consumer indifferent between the two products is given by the taste \( \bar{\omega} = v_W - p - v_L \). Demand for \( W \) as a function of the price is therefore the integral of all consumers with taste \( \omega \in [-t/2, \bar{\omega}] \). Likewise, demand for \( L \) as a function of the price of \( W \) is integral of consumers with taste \( \omega \in (\bar{\omega}, t/2] \). The subscript "0" in the demand functions below denotes absence of network externalities.

\[
W^\text{new}_0(p) = \begin{cases} 0 & \text{if } v_W - p - v_L \leq -\frac{t}{2} \\ \sigma \left( \frac{1}{2} + \frac{v_W - v_L - p}{t} \right) & \text{if } -\frac{t}{2} \leq v_W - p - v_L \leq \frac{t}{2} \\ \sigma & \text{if } v_W - p - v_L \geq \frac{t}{2} \end{cases} 
\]

\[
L^\text{new}_0(p) = \begin{cases} \sigma & \text{if } v_W - p - v_L \leq -\frac{t}{2} \\ \sigma \left( \frac{1}{2} - \frac{v_W - v_L - p}{t} \right) & \text{if } -\frac{t}{2} \leq v_W - p - v_L \leq \frac{t}{2} \\ 0 & \text{if } v_W - p - v_L \geq \frac{t}{2} \end{cases} 
\]

I will only focus on interior solutions for the study of the market without network externalities, i.e. \( t \) is assumed to be sufficiently large in order for the restriction \(-\frac{t}{2} \leq v_W - p - v_L \leq \frac{t}{2}\) to be true for any profit maximizing price, \( p \).\footnote{It will be clear shortly that this restriction specifically requires \( \frac{v_W - v_L}{t} \in [-\frac{1}{2}, \frac{1}{2}] \).} Figure 1 provides a graphical representation of the market shares at a given price.

Maintaining the assumption on interior solution, the profit maximizing price for the developer of \( W \) is straightforward.

\[ p_0 = \arg \max_p W^\text{new}_0(p)p \]

The first order condition solves the optimal price:

\[ p_0 = \frac{t}{4} + \frac{1}{2} (v_W - v_L) \quad (3) \]

By inserting equation (3) into equations (1) and (2), equilibrium user shares for the new user population are found:

\[ W^\text{new}_0(p_0) = \sigma \left( \frac{1}{4} + \frac{v_W - v_L}{2t} \right) \]

\[ L^\text{new}_0(p_0) = \frac{t}{4} + \frac{1}{2} (v_W - v_L) \]
\[ L_0^{new}(p_0) = \sigma \left( \frac{3}{4} - \frac{v_W - v_L}{2t} \right) \]

From the above expressions it is easily seen that, in the case of equal quality for the median consumer, the market will be divided with 1/4 of the new users users purchasing \( W \), while the rest 3/4 will download \( L \) for free. This follows from the assumption of zero marginal cost and linear demand functions.\(^{10}\)

The profits to the developer of \( W \) at optimal price \( p_0 \) is:

\[ \pi(p_0) = \sigma \left[ \frac{1}{4} + \frac{v_W - v_L}{2t} \right][\frac{t}{4} + \frac{1}{2}(v_W - v_L)] \] (4)

The \textit{total} user shares between the two software products as functions of price \( p \), i.e. the sum of old and new users, are given by the following equations:

\[ W_0(p) = \sigma \left( \frac{1}{2} + \frac{v_W - v_L - p}{t} \right) + (1 - \sigma) \frac{1 + \lambda}{2} \] (5)

\[ L_0(p) = \sigma \left( \frac{1}{2} - \frac{v_W - v_L - p}{t} \right) + (1 - \sigma) \frac{1 - \lambda}{2}. \] (6)

The total installed base for each software type is not of importance when network externalities are absent. In the next subsection equations (5) and (6) turn out to be important when the developer of \( W \) considers the optimal pricing strategy.

\subsection*{2.2 Network externalities}

When modeling network externalities, the model assumes the existence of a \textit{critical mass}, a minimum number of adopters of a technology sufficient for self-sustained further adoption, or following the definition of Economides and Himmelberg (1995): a critical mass is the smallest network supported in an equilibrium. Alternatively, the critical mass can be interpreted as a \textit{tipping point} in which the adoption converges to zero for a network size below the tipping point, while adoption converges to some positive number larger than the tipping point otherwise. As pointed out by among others Allen (1988), the outcome of such markets is difficult to predict since a prospective member of a network will decide to join the network only if it is expected that at least a critical mass of others also join. Hence, markets with the existence of a critical mass typically have multiple equilibria based on whether consumers are successfully able to coordinate their actions.

\(^{10}\)A similar model in Lerner and Schankerman (2010) gives the same result.
Such game theoretical considerations are avoided by implicitly assuming expectations based on market shares without network externalities. If e.g. a prospective user observes that total user share for \( L \) in equation (6) is above the critical mass at a given price when network externalities are absent, she expects the situation to persist in the presence of network externalities. On the other hand, if this user share is below the critical mass, she will expect that no one will obtain \( L \), and will therefore choose not to download \( L \) herself.

Network externalities are added to the model by introducing the parameter \( N \in [0, 1/2] \) which represents the mentioned critical mass demanded by new users to be willing to obtain a piece of software in question. Technically, the model assumes that users’ valuation is zero for a piece of software with user share below critical mass. I denote the median new user’s valuation for the two software goods as \( u_W \) and \( u_L \). A new user at location \( \omega \) will value the two products as follows:

\[
\begin{align*}
    u_W - \frac{1}{2} \omega &= \begin{cases} 
    v_W - \frac{1}{2} \omega & \text{if } W_0(p) > N \\
    0 & \text{if } W_0(p) \leq N
    \end{cases} \\
    u_L + \frac{1}{2} \omega &= \begin{cases} 
    v_L + \frac{1}{2} \omega & \text{if } L_0(p) > N \\
    0 & \text{if } L_0(p) \leq N
    \end{cases}
\end{align*}
\]

Hence, if the user share of a software good is above critical mass, the perceived quality is the same as in section 2.1 where there are no network externalities involved. Otherwise, the software good in question is perceived as worthless.

Demand functions for the two software types, depending on user shares being above critical mass, can be shown to be:

\[
\begin{align*}
    W_N^{new}(p) &= \begin{cases} 
    W_0^{new}(p) & \text{if } W_0 > N \text{ and } L_0 > N \\
    \min\{\sigma, \sigma\left(\frac{1}{2} + \frac{v_W - p}{t}\right)\} & \text{if } W_0 > N \text{ and } L_0 \leq N \\
    0 & \text{if } W_0 \leq N \text{ and } L_0 > N
    \end{cases} \\
    L_N^{new}(p) &= \begin{cases} 
    L_0^{new}(p) & \text{if } W_0 > N \text{ and } L_0 > N \\
    \min\{\sigma, \sigma\left(\frac{1}{2} + \frac{v_L - p}{t}\right)\} & \text{if } W_0 \leq N \text{ and } L_0 > N
    \end{cases}
\end{align*}
\]

The subscript "\( N \)" in the demand functions denotes the presence of network externalities.\(^{11}\) \( W_0 \) and \( L_0 \) are described in equations (5) and (6). Hence, if demand for

\(^{11}\)The subscript "\( 0 \)" can therefore be interpreted as a special case where \( N = 0 \).
Figure 2: Network strategic pricing will lead to increased demand for $W$ following lack of viable open source substitutes.

both software types exceeds the critical mass, then the demand functions are identical to those in the benchmark where network externalities are unimportant. Provided that $L^{old} < N$, the developer of the commercial software type has an additional pricing strategy: To set the price such that $L_N^{new}(p) = 0$. From equations (7) and (8) it is clear that this pricing strategy will lead to a positive shift in demand for $W$, while demand for $L$ will be zero among new users. The effect on demand for $W$ from network strategic pricing is illustrated in Figure 2: For prices less than or equal to $p_N$, the user share of $L$ is below critical mass and is regarded as worthless. Therefore, a discontinuous increase in demand for $W$ occurs at this exact price due to lack of valuable substitutes. In effect, $W$ is a monopolist when $L_0(p) \leq N$. However, the monopoly power is constrained since $W$ is limited in the price setting since any price above the threshold such that $L_0(p) > N$ would cause $W$ to lose its monopoly position.

For the remainder of the paper I will assume that parameter values are such that $W_0(p_0) > N$ and $L_0(p_0) > N$.\footnote{In other words, $N - L^{old} < \sigma \left( \frac{3}{4} - \frac{v_W - v_L}{2t} \right)$, and $N - W^{old} < \sigma \left( \frac{1}{4} + \frac{v_W - v_L}{2t} \right)$.} That is, at the profit maximizing prices where network externalities are absent, both software types have user shares above critical mass. Hence, the developer of $W$ does not need to lower the price to ensure its own existence, and user shares of $L$ will not fall below critical mass unless the seller of $W$ makes a strategic pricing decision with the attempt to pull demand for $L$ below critical mass.

The network strategic price, $p_N$, is given by the equality $L_0^{new}(p_N) + L^{old} = N$. The
Figure 3: Profit functions with kinks. Network strategic pricing more likely for larger $N$.
Parameter values: $\sigma = 0.5, v_W = v_L = t = \lambda = 1$.

The solution is:

$$p_N = \frac{t}{\sigma} (N - L_{old}) + v_W - v_L - \frac{t}{2}.$$  \hspace{2cm} (9)

As before, $L_{old} = (1 - \sigma)(1 - \lambda)/2$. Conditional on the network strategic price being imposed, the profits for the seller of $W$ is given by:

$$\pi(p_N) = \min \{ \sigma, \sigma(1 + \frac{v_L}{t}) - N + L_{old} \} \left( \frac{t}{\sigma} (N - L_{old}) + v_W - v_L - \frac{t}{2} \right).$$ \hspace{2cm} (10)

Obviously, the output effect by imposing network strategic pricing is large. However network strategic pricing may also imply a large price reduction. As is clear from equation (9), $p_N$ is increasing in $N$. Intuitively, the larger the critical mass is, the less the price must be reduced in order to force demand for $L$ below critical mass.

As Figure 3 illustrates, the size of $N$ is crucial for the optimal pricing strategy. The profit function will have a kink at the network strategic price, $p_N$: For $p \leq p_N$, new users will not demand $L$ for any price, while for $p > p_N$, both software types exist in the market for new users, and demand for $W$ will drop once the price exceed the threshold $p_N$. The larger $N$ is, the smaller is the price differential between $p_N$ and $p_0$, and $p_N$ is more likely to be the profit maximizing price.

Among other parameters of interest it is easy to check that $\frac{\partial \pi(p_N)}{\partial \lambda} > \frac{\partial \pi(p_{0})}{\partial \lambda} = 0$, and $\frac{\partial \pi(p_N)}{\partial t} < \frac{\partial \pi(p_{0})}{\partial t}$ for sufficiently small $N$.\textsuperscript{13}

\textsuperscript{13}If $N - L_{old} \leq \frac{\sigma v_L}{t}$, then the restriction is $N - L_{old} < \sigma \left[ \frac{9}{16} - \frac{(v_W - v_L)^2}{4t^2} \right]$. 

13
Hence, in the absence of piracy, network strategic pricing is more likely to occur whenever $W$ has the benefit of a large installed base ($\lambda$ close to one) and when the degree of differentiation is small (small $t$).

### 2.3 Welfare

Since long-run considerations from piracy and pricing strategies in OSS vs. commercial software competition is not considered in the model, a welfare analysis is somewhat limited. Within the model one unambiguous welfare result is derived: the social optimum will not be realized in the market solution. When network externalities are either not present or not important for the pricing decision, too few will use the commercial software type and too many will use open source software. When network strategic pricing is imposed the opposite is the case: too many will use commercial software and too few (none) will use open source software.

#### 2.3.1 First best

Since marginal cost is assumed to be zero and investment costs are sunk, the social surplus is equivalent to the gross consumer surplus. The payoff of the old users is not defined, thus the welfare analysis will only apply to the gross welfare of new users. Accordingly, the welfare results will be normalized to $\sigma$, the new user population size. The gross value of using $W$ and $L$ for a consumer at location $\omega$ is $v_W - \frac{1}{2} \omega$ and $v_L + \frac{1}{2} \omega$, respectively. First best allocation is given by:

$$\max_m \int_{-t/2}^m (v_W - \frac{1}{2} \omega)d\omega + \int_{m}^{t/2} (v_L + \frac{1}{2} \omega)d\omega,$$

where $m$ is the marginal user who values the two software products equally. The maximization problem has the solution:

$$m^* = v_W - v_L.$$

The first-best user shares among new users are then given by:

$$W^* = \frac{1}{2} + \frac{v_W - v_L}{t}$$

$$L^* = \frac{1}{2} - \frac{v_W - v_L}{t}.$$

If $N - L^{old} > \frac{\sigma v_L}{t}$, then the restriction is $N - L^{old} < \sigma \frac{3}{4} + \frac{1}{2t} \sqrt{(v_L - v_W)(5v_L - v_W)}.$
The maximized social surplus becomes:

\[ W F^{FB} = \frac{t^2}{8} + \frac{t}{2} (v_W + v_L) + \frac{1}{2} (v_W - v_L)^2. \]

First best is only implemented if the commercial software providers gives away the product for free. Arguably, the first best allocation is unlikely to be achieved. The rest of the welfare analysis will therefore focus on welfare in the market solution.

2.3.2 No network externalities

When network externalities are non-existent or is not accounted for in the pricing decision, the profit maximizing price is \( p_0 \) given by equation (3). It can be shown that under this price the consumer indifferent between \( W \) and \( L \) will be located at:

\[ m_0 = \frac{-t}{4} + \frac{1}{2} (v_W - v_L) = \frac{-t}{4} + \frac{1}{2} m^*. \]

The social surplus under price \( p_0 \) is then:

\[ W F_0 = W F^{FB} - \frac{t}{4} (v_W - v_L) - \frac{1}{8} (v_W - v_L - t/2)^2 \]

Welfare result 1. The user share of commercial software is too small, and user share of open source software is too large in the market solution with price \( p_0 \).

The reason for the welfare maximum not being attained is due to the market power the commercial software provider obtains from product differentiation. Marginal cost is zero, and the price is higher than the marginal cost at the profit maximizing price. Hence, some consumers will choose to download the open-source software product for free, even if the gross benefit from the commercial software type is higher.

2.3.3 Network strategic pricing

When network strategic pricing is imposed there are two opposing effects in play. On one hand, open source software will not be available which will reduce the payoff of individuals with high values of \( \omega \). On the other hand the commercial software product will be available at a lower price. The location of the worst-off consumers who will obtain \( W \) is:

\[ m_N = t/2 - \min \{0, v_L - t(N - L^{old})/\sigma\}. \]
Note that if \((N - L^{old})t < \sigma v_L\) there will be no effects on aggregate consumption. The social surplus when network strategic pricing is given by:

\[
WF_N = \int_{-t/2}^{m_N} (v_W - \frac{1}{2} \omega) d\omega
\]

\[
= \begin{cases} 
  tv_W & \text{if } (N - L^{old})t \leq \sigma v_L \\
  tv_W - \frac{1}{2}(v_W - \frac{t}{8}) - (\frac{N-L^{old}}{\sigma} - v_L)(v_W - \frac{1}{4}(\frac{N-L^{old}}{\sigma})) & \text{if } (N - L^{old})t > \sigma v_L 
\end{cases}
\]

Intuitively, the degree of differentiation matters, even if there are no aggregate effects on quantity. If the degree of differentiation is small, the negative welfare effects are not substantial since \(W\) is a close substitute to \(L\). If the two products are more differentiated, welfare is more likely to be reduced. The following welfare result summarize the welfare effects relative to the price where network externalities are not directly accounted for in the price.

**Welfare result 2.** *Network strategic pricing increases welfare relative to \(p_0\) if and only if \(v_W > v_L\) and:*

\[
t < 6(v_W - v_L).
\]

This welfare result implies that network strategic pricing will always reduce welfare if there are no quality differences between the two products for the median user, \(v_W = v_L\) since \(t\) is strictly positive. When there are no quality differences, the price \(p_0\) yields user shares of 1/4 and 3/4 between \(W\) and \(L\). When network strategic pricing is imposed a share 1/4 will increase their gross surplus (user with \(\omega < 0\) who download \(L\) at price \(p_0\)), while a share 1/2 will be worse off (all users with \(\omega > 0\)).

Relative to the social optimum there will be over-consumption of commercial software and under-consumption (zero) of open source software.

**Welfare result 3.** *The user share of commercial software is too large, and user share of open source software is too small in the market solution with price \(p_N\).*

### 3 Piracy

#### 3.1 No network externalities

Piracy is added to the model by allowing users to obtain a free-of-charge version of \(W\). Despite \(W\) being available for free, there may be reasons for users rather to purchase
W legally. First, a pirated version may have some degree of quality degradation. This may for instance be due to lack of customer support or incompatibility with patches.\footnote{A patch is a software update designed to fix bugs and security vulnerabilities.} Second, users may face constraints, for both technical and ethical reasons, of obtaining illegal copies of a piece of software. Moreover, such non-monetary costs may vary between individuals.

The timing of the choices of the new users is illustrated in Figure 4 and is as follows: (i) Users choose between joining the $W$ network and joining the $L$ network. (ii) Users in the $W$ network choose whether to obtain a pirated or a legal version. In addition it is assumed that new users are unaware of their personal costs of obtaining pirated software before choosing network in stage (i). Instead, they will make their choices based on the expected value of choosing a software type at given price.\footnote{Solving the model with full initial information on personal “cost of copying” is feasible, but will in addition to the distributions of $\omega$ and $\eta$, depend on the distribution of $\omega + \eta$ which greatly complicates model since this distribution generally will not be continuous over the area of interest.} Not being aware of one’s personal cost of copying is a fairly strong assumption, but can be partly justified with regard to swift changes in technology and legislation in this specific area. Since consumers not necessarily are updated on how to by-pass the latest type of copy protection, or what the consequences are of violating IP laws according to the latest legislation,\footnote{When obtaining a pirated software product, one will typically get a version where the copy protection already is ”cracked” by a third party. However some effort is typically required by the end-user in order for the cracked version to function. Such efforts may involve, e.g., replacing certain files in the installed directory, and methods to mimic online activation of the product.} the information of one’s personal cost-of-copying may be revealed only after one has decided to obtain a proprietary software product.

Legal and illegal versions of $W$ differ for new users in two ways: The pirated version is degraded in quality by a parameter $\alpha \in [0, 1]$, where $\alpha = 0$ means that there is no difference in quality between legal and illegal copies, and $\alpha = 1$ means that the pirated version is of no value. Also, new users differ by the personal cost, $\eta$, of obtaining pirated copies which is uniformly distributed over the interval $[0, z]$. This parameter is meant to capture the above mentioned observation that individuals differ in their technical abilities and ethical constraints in obtaining pirated software.

Demand functions are found by solving the model backwards. Given that a user has chosen to join the $W$ network, her value of $\eta$ is revealed. A new user will choose to
purchase rather than pirate whenever the payoff is higher for purchasing:
\[ v_W - p \geq (1 - \alpha)v_W - \eta. \]

Otherwise, piracy is chosen. From the uniform distribution of \( \eta \) it is straightforward to calculate the share (or probability) of new users purchasing \( W \):
\[ \Pr(\text{buy}) = \phi(p) = \min\{1 - \frac{p - \alpha v_W}{z}, 1\}. \tag{11} \]

Hence, the piracy rate among new users of \( W \), i.e. the share who do not purchase is:
\[ \Pr(\text{pirate}) = 1 - \phi(p) = \max\{\frac{p - \alpha v_W}{z}, 0\} \tag{12} \]

Intuitively, the piracy rate increases in the purchase price. Moreover, \( z \) is a measure of the price-sensitivity of the piracy rate: A unit increase in the purchasing price will increase the piracy rate by \( 1/z \) percentage points whenever \( p > \alpha v_W \).

By backward induction, a new user will join the \( W \) network if the expected value of choosing \( W \) is larger than the value of downloading \( L \). In the first stage a new user will either purchase or pirate \( W \) rather than downloading \( L \) whenever:
\[ \phi(p)[v_W - p] + (1 - \phi(p))[(1 - \alpha)v_W - \frac{(p - \alpha v_W)}{2}] - v_L \geq \omega, \]
where \( \phi(p) \) and \( 1 - \phi(p) \) are given in (11) and (12), and \( (p - \alpha v_W)/2 \) is the expected value of \( \eta \) conditional on being a pirate.\footnote{\( E[\eta|\eta \leq p - \alpha v_W] = \frac{p - \alpha v_W}{2} \).} This implies that a new user will download
when the opposite is true, i.e.
\[ \phi(p)[v_W - p] + (1 - \phi(p))[(1 - \alpha)v_W - (p - \alpha v_W)/2] - v_L < \omega. \]

The demand functions for the two software products among new users, independent on pirating or purchasing, are found by integrating over all consumers who satisfy the above inequalities.\(^\text{18}\)

\[
\tilde{W}_{0}^{\text{new}}(p) = \sigma \left[ \frac{1}{2} + \frac{\phi(p)[v_W - p] + (1 - \phi(p))[(1 - \alpha)v_W - (p - \alpha v_W)/2] - v_L}{t} \right]
\]
\[
\tilde{L}_{0}^{\text{new}}(p) = \sigma \left[ \frac{1}{2} - \frac{\phi(p)[v_W - p] + (1 - \phi(p))[(1 - \alpha)v_W - (p - \alpha v_W)/2] - v_L}{t} \right]
\]

From (13) and (14), and their equivalents in (1) and (2), the following result is of interest:

**Proposition 1.** At a given price, total demand for \(W\) is larger with piracy than without piracy. Equivalently, total demand for \(L\) is smaller with piracy than without piracy:

\[
\tilde{W}_{0}^{\text{new}}(p) > W_{0}^{\text{new}}(p) \quad \text{and} \quad \tilde{L}_{0}^{\text{new}}(p) < L_{0}^{\text{new}}(p) \quad \text{if} \quad \phi(p) < 1.
\]

Proof. The inequalities \(\tilde{W}_{0}^{\text{new}}(p) > W_{0}^{\text{new}}(p)\) and \(\tilde{L}_{0}^{\text{new}}(p) < L_{0}^{\text{new}}(p)\) is equivalent to:

\[
\frac{1}{2}[p(1 + \phi(p)) + \alpha v_W(1 - \phi(p))] < p
\]

\[\Leftrightarrow \frac{-(p - \alpha v_W)^2}{2z} < 0
\]

\[\Leftrightarrow p > \alpha v_W.
\]

This is also the condition for a positive piracy rate, i.e. \(\phi(p) < 1\). \(\square\)

The demand function for purchased (pirated) versions of \(W\) among new users is the probability of both being a \(W\) user and a legal user (pirate). Hence, by rearranging (13) and multiplying with the probabilities of purchasing and pirating, respectively, the demand functions turn out to be:

\[
\tilde{W}_{\text{buy},0}^{\text{new}}(p) = \phi(p)\sigma \left[ \frac{1}{2} + \frac{v_W - \frac{1}{2}[p(1 + \phi(p)) + \alpha v_W(1 - \phi(p))] - v_L}{t} \right]
\]

\(^{18}\)A diacritical tilde is included on relevant functions and results to distinguish the piracy case from the no-piracy benchmark.
\[ \tilde{W}^{\text{new},0}_{\text{piracy}}(p) = (1 - \phi(p))\sigma \left[ \frac{1}{2} + \frac{v_{W} - \frac{1}{2}p(1 + \phi(p)) + \alpha v_{W}(1 - \phi(p))}{t} - v_L \right]. \] (16)

The profit maximizing price for the developer of \( W \) is then:

\[ \tilde{p}_0 = \arg \max_p W^{\text{new}}_{\text{buy},0}(p). \] (17)

Hence, \( \tilde{p}_0 = p^* \) is determined by the first order condition of the above maximization problem:

\[ \tilde{W}^{\text{new}}_{\text{buy},0}(p^*) + p^*[\phi'(p^*)\tilde{W}^{\text{new}}_0(p^*) + \phi(p^*)\frac{\partial \tilde{W}^{\text{new}}_0(p)}{\partial p}] = 0. \] (18)

Here, \( \tilde{W}^{\text{new}}_0(p) \) is given by equation (13). From equation (18), it can be seen that the negative output effect from a marginal price increase can be broken down into two parts: \( \phi(p)(\partial \tilde{W}^{\text{new}}_0(p)/\partial p) \) is the decrease in demand for purchased \( W \) in favor of downloading \( L \), while \( \phi'(p)\tilde{W}^{\text{new}}_0(p) \) represents the loss in revenue from an increase in the piracy rate.\(^19\)

**Proposition 2.** The price with piracy can be both higher and lower than without piracy, i.e., both \( \tilde{p}_0 > p_0 \) and \( \tilde{p}_0 \leq p_0 \) is feasible.

**Proof.** Appendix. \( \square \)

It is not difficult to find instances where the price with piracy, \( \tilde{p}_0 \), is lower than the price without, \( p_0 \). The opposite case is true only for a small subset of parameters, which is proven in the appendix. Nevertheless, the result may seem like a paradox as one would normally expect the competitive pressure from software piracy to always lead to a price reduction. \( \tilde{p}_0 \) exceeding \( p_0 \) can only be true if the degree of degradation of illegal copies, \( \alpha \), is close to or equals zero, i.e. that legal and illegal copies are more or less perfect substitutes. For \( \alpha = 0 \), the *expected* cost of obtaining any copy of \( W \) is \( p(1 + \phi(p))/2 < p \) for \( \phi(p) < 1 \). Thus setting an ex-post price \( \tilde{p}_0 \) slightly above \( p_0 \) may be optimal, since the ex-ante cost of \( W \) may still be below \( p_0 \). Note, however, that the mark-up above \( p_0 \), \((\tilde{p}_0 - p_0)/p_0\), will be small,\(^20\) and it can hardly be argued that \( \tilde{p}_0 > \tilde{p} \) is a likely outcome within the model setup. Nonetheless, the result gives a rationale for why \( \tilde{p}_0 \) may not necessarily be monotonically increasing in \( \alpha \).

\(^{19}\)The solution to (18) is the only non-complex solution of a cubic equation, and is omitted.

\(^{20}\)From numerical computations I did not encounter a markup exceeding 2%.
Another observation of the optimal price with piracy, which turns out to be of importance when (re-)extending the model to include network externalities, is summarized by the following proposition:

**Proposition 3.** The price will eventually seize to increase in the degree of differentiation:

\[
\frac{dp_0}{dt} > 0, \quad \lim_{t \to \infty} \frac{dp_0}{dt} = 0.
\]

**Proof.** Write the first-order condition for profit maximization as:

\[
[\phi(p^*) + p^* \phi'(p^*)]f(p^*, t) + p^* \phi(p^*) f_p'(p^*, t) = 0,
\]

where \( f(p^*, t) = \frac{1}{2} + \frac{v_W - \frac{1}{2}[p(1 + \phi(p)) + \alpha v_W (1 - \phi(p))] - v_L}{t} \). By applying implicit differentiation w.r.t. \( p^* \) and \( t \), the following expression is found:

\[
\frac{dp^*}{dt} = -\frac{[\phi(p^*) + p^* \phi'(p^*)]f_t' + p^* \phi(p^*) f_p''}{2\phi'(p^*)f + 2[\phi(p^*) + p^* \phi'(p^*)]f_p' + p^* \phi(p^*) f_p''}
\]

It is easy to check that \( \lim_{t \to \infty} f_t'(p^*, t) = 0, \lim_{t \to \infty} f_p'(p^*, t) = 0, \lim_{t \to \infty} f_p''(p^*, t) = 0, f_p'(p^*, t) = 0 \) and \( \lim_{t \to \infty} f(p^*, t) = 1/2 \). Hence,

\[
\lim_{t \to \infty} \frac{dp^*}{dt} = \frac{0}{\phi'(p^*)} = 0.
\]

The intuition behind Proposition 3 is as follows: As \( t \) increases, a larger share of the population will have strong preferences for \( W \) which leads to an increase in price. On the other hand, as their preferences towards piracy are unchanged, there is a limit to how high the price can be set without the \( W \) developer losing all the customers to pirated versions of the product.

Table 1 provides a few examples on how prices, and therefore other endogenous variables, are affected by changes in parameter values. In all the examples the qualities of the commercial and open source products are equal. The price does increase for an increase in the degree of product differentiation \( t \) when piracy is present, but at a smaller scale relative to the benchmark case without piracy, which is consistent with proposition 3.

21
Higher degradation in quality of pirated versions, i.e., an increase in $\alpha$, has an ambiguous effect on the price. On one hand, piracy is now a less valuable option since the competitive pressure from piracy will be weaker, which allows for higher prices. On the other hand, a higher degree of degradation may lower the *ex-ante* value of obtaining $W$. Hence, for low values of $\alpha$ the price may decrease in $\alpha$ since the piracy rate initially is non-negligible, and the value of obtaining any version of $W$ may fall. In a small subset of parameter values the price may even exceed $p_0$ when $\alpha$ is decreasing (from proposition 2). In the numerical examples in table 1 the latter effect dominates which explains why the price increases when the quality of illegal copies is improved.

Lastly, a decrease in $z$ lowers the expected cost of copying resulting in a more price sensitive piracy rate. When individuals are more likely to acquire illegal copies at a given price, the seller of $W$ responds by cutting the price, which ultimately leads to a larger $W$ network and a smaller $L$ network. However, due to the result in Proposition

<table>
<thead>
<tr>
<th>Parameter Values</th>
<th>Price</th>
<th>No piracy</th>
<th>Piracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma = 1$</td>
<td>$v_W = v_L = 1$</td>
<td>price</td>
<td>0.25</td>
</tr>
<tr>
<td>$t = 1$</td>
<td>$W^{new}$ network</td>
<td>0.25</td>
<td>0.312</td>
</tr>
<tr>
<td>$\alpha = 0.1$</td>
<td>$L^{new}$ network</td>
<td>0.75</td>
<td>0.688</td>
</tr>
<tr>
<td>$z = 0.5$</td>
<td>piracy rate</td>
<td>0%</td>
<td>19.5%</td>
</tr>
<tr>
<td></td>
<td>profits</td>
<td>0.0625</td>
<td>0.049</td>
</tr>
<tr>
<td>$\sigma = 1$</td>
<td>$v_W = v_L = 1$</td>
<td>price</td>
<td>0.5</td>
</tr>
<tr>
<td>$t = 2$</td>
<td>$W^{new}$ network</td>
<td>0.25</td>
<td>0.38</td>
</tr>
<tr>
<td>$\alpha = 0.1$</td>
<td>$L^{new}$ network</td>
<td>0.75</td>
<td>0.62</td>
</tr>
<tr>
<td>$z = 0.5$</td>
<td>piracy rate</td>
<td>0%</td>
<td>32%</td>
</tr>
<tr>
<td></td>
<td>profits</td>
<td>0.125</td>
<td>0.068</td>
</tr>
<tr>
<td>$\sigma = 1$</td>
<td>$v_W = v_L = 1$</td>
<td>price</td>
<td>0.25</td>
</tr>
<tr>
<td>$t = 1$</td>
<td>$W^{new}$ network</td>
<td>0.25</td>
<td>0.32</td>
</tr>
<tr>
<td>$\alpha = 0.05$</td>
<td>$L^{new}$ network</td>
<td>0.75</td>
<td>0.68</td>
</tr>
<tr>
<td>$z = 0.5$</td>
<td>piracy rate</td>
<td>0%</td>
<td>29.8%</td>
</tr>
<tr>
<td></td>
<td>profits</td>
<td>0.0625</td>
<td>0.045</td>
</tr>
<tr>
<td>$\sigma = 1$</td>
<td>$v_W = v_L = 1$</td>
<td>price</td>
<td>0.25</td>
</tr>
<tr>
<td>$t = 1$</td>
<td>$W^{new}$ network</td>
<td>0.25</td>
<td>0.35</td>
</tr>
<tr>
<td>$\alpha = 0.1$</td>
<td>$L^{new}$ network</td>
<td>0.75</td>
<td>0.65</td>
</tr>
<tr>
<td>$z = 0.3$</td>
<td>piracy rate</td>
<td>0%</td>
<td>18.5%</td>
</tr>
<tr>
<td></td>
<td>profits</td>
<td>0.0625</td>
<td>0.44</td>
</tr>
</tbody>
</table>

Table 1: Price, network size, piracy rate and profits for various parameter values without network externalities. Values in bold are changes relative to the first row.
2 there may exist a small interval where \( p \) is decreasing in \( z \) by a small amount for very low values of \( \alpha \).

### 3.2 Network externalities

In the model with both piracy and network externalities, the seller of the commercial software product will now have the option to impose such strategic pricing. However, as a pirated version of \( W \) may exist, pirates will contribute to the network size of \( W \) and thus alter the incentives for network strategic pricing. It turns out that network strategic pricing is more likely to be optimal in the presence of piracy than without it. Piracy increases the total user share of \( W \) and, with the exception of a small subset of parameters, decreases the price of legal versions. As a result the price reduction by imposing network strategic pricing is smaller. Hence, the cost of introducing a price where the entire market is captured is reduced, and is more likely to be the optimal strategy.

As before, I assume that users’ expectations on market shares at given prices are based on the hypothetical market shares in the absence of network externalities. Moreover, at the profit-maximizing price \( \tilde{p}_0 \) given by equation (17), it is assumed that parameters are such that \( \tilde{W}_0^{new}(\tilde{p}_0) + \tilde{W}^{old} > N \) and \( \tilde{L}_0^{new}(\tilde{p}_0) + \tilde{L}^{old} > N \). I.e., for the profit maximizing price of \( W \) in the absence of network externalities, user shares of both software types exceed the critical mass needed to generate demand among new users.

The network strategic price with piracy is equivalent to its counterpart in the no-piracy benchmark. The network strategic price, \( \tilde{p}_N \), is the highest price of \( W \) in which demand for \( L \) is below the critical mass, i.e., the price is set such that \( \tilde{L}_0^{new}(\tilde{p}_N) + \tilde{L}^{old} = N \). In other words, \( \tilde{p}_N \) solves:

\[
\sigma \left[ \frac{1}{2} - \frac{W - \frac{1}{2}[\tilde{p}_N(1 + \phi(\tilde{p}_N))] + \alpha v_W(1 - \phi(\tilde{p}_N))] - v_L}{t} \right] + L_{old} = N. \tag{19}
\]

If \( \phi(\tilde{p}_N) = 1 \), i.e. \( \tilde{p}_N < \alpha v_W \), the solution to the above expression is equal to the network strategic price, \( p_N \), in equation (9) from the no-piracy benchmark. For \( \phi(\tilde{p}_N) < 1 \), i.e. \( \tilde{p}_N > \alpha v_W \), the solution to the above expression is:

\[
\tilde{p}_N = \alpha v_W + z - \sqrt{z \left[ 2t \left( \frac{1}{2} - \frac{N - L^{old}}{\sigma} \right) + v_L - (1 - \alpha)v_W \right].}
\]

Or, inserting \( p_N \) from equation (9):
\[ \tilde{p}_N = \alpha v_w + z - \sqrt{z \left[ z - 2(p_N - \alpha v_w) \right]} . \] (20)

Hence, there is a non-linear relationship between the network strategic prices when piracy is present and in the no-piracy benchmark. From (20) it can be shown that the network strategic price in the no-piracy benchmark can never exceed its equivalent in the presence of piracy.

**Proposition 4.** The network strategic price with piracy is higher than or equal to the one without piracy: \( \tilde{p}_N \geq p_N \).

**Proof.** For \( \tilde{p}_N \leq \alpha v_w \) the piracy rate is zero, i.e. \( \phi(\tilde{p}_N) = 1 \Rightarrow \tilde{L}_0^{new}(\tilde{p}_N) = L_0^{new}(\tilde{p}_N) \), which in turn implies that \( \tilde{p}_N = p_N \).

For \( \tilde{p}_N > \alpha v_w \), the piracy rate will be positive. Solving equation (20) for \( p_N \) yields the following expression:

\[ p_N = \tilde{p}_N - \frac{1}{2z}(\tilde{p}_N - \alpha v_w)^2 , \]

which clearly implies that \( \tilde{p}_N > p_N \) for \( \tilde{p}_N > \alpha v_w \). \( \square \)

This result can be intuitively explained by the fact that \( \tilde{L}_0(p) \leq L_0(p) \), by Proposition 1. By definition, \( L_0(p_N) = N \), which implies the inequality \( \tilde{L}_0(p_N) \leq N \). Hence, for the same \( N \), the price can be increased from \( p_N \) while still maintaining demand for \( L \) below critical mass.

The network strategic price, \( \tilde{p}_N \), reacts similarly to changes in parameter values to its counterpart in the no-piracy benchmark. Obviously, stronger network externalities, interpreted as a larger critical mass \( N \), and a larger installed base of \( W \) users (\( \lambda \)) lead to an increase in \( \tilde{p}_N \). The response from increased degree of differentiation, \( t \), depends on the strength of network externalities, is summarized by the following proposition:

**Proposition 5.** If \( N - L^{old} > \sigma/2 \), then \( \tilde{p}_N \) increases convexly in \( t \): \( \frac{dp_N}{dt} > 0 \) and \( \frac{d^2p_N}{dt^2} > 0 \).

**Proof.** The proof follows directly by applying implicit differentiation w.r.t. \( t \) in equation (19):

\[ \frac{dp}{dt} = \frac{N - L^{old}}{\phi(p)} - \frac{1}{2} > 0 \text{ if } N - L^{old} > \sigma/2 , \]

\[ \frac{d^2p}{dt^2} = \frac{N - L^{old}}{z\phi(p)^2} - \frac{1}{2} \frac{dp}{dt} > 0 \text{ if } \phi(p) < 1 . \]

See appendix for an alternative proof. \( \square \)
\[ \tilde{p}_N \text{ is increasing in } t \text{ for sufficiently strong network externalities. The intuition behind this result is that a higher degree of differentiation yields lower demand for } L \text{ at any price } p > \bar{p}, \text{ where } \bar{p} \text{ is the price that solves } L_{0}^{\text{new}}(\bar{p}) = \sigma/2. \text{ Hence, the price reduction from } \tilde{p}_0 \text{ that network strategic pricing implies will be lower for large } t. \text{ Moreover, following the concavity of } L_{0}^{\text{new}}(p) \text{ in } p, \partial^2 \tilde{p}_N / \partial t^2 > 0 \text{ whenever a solution exists. A graphical illustration of the relationship between } \tilde{p}_N \text{ and } t \text{ is provided in Figure 5.}

Comparative statics is very different in cases with and without network externalities, when piracy is present. While \( \tilde{p}_0 \) converges to a constant in \( t \), \( \tilde{p}_N \) increases convexly in \( t \) for \( \tilde{p}_N < \tilde{p}_0 \), which is always true given the assumptions in the model.\(^{21}\)

3.2.1 Special case: Limit pricing

Following Bae and Choi (2006) limit pricing with piracy is defined by the price in which no user has incentives to obtain pirated products. From equation (12) it is clear that the piracy rate is zero whenever \( p \leq \alpha v W \).

When the piracy rate is zero, network strategic pricing is given by equation (9), and it can be shown that if

\[
\frac{t}{\sigma}(N - L^{old}) + v_W(1 - \alpha) - v_L - \frac{t}{2} < 0,
\]

then the network strategic price is indeed below the no-piracy limit price. In the numer-

\(^{21}\) \( \tilde{p}_N < \tilde{p}_0 \) is true whenever \( \tilde{L}_0(\tilde{p}_0) > N \) which holds by assumption.
ical analysis in section 4, it will be shown that it may be a profit-maximizing strategy to set the price below the no-piracy limit price.

3.3 Welfare

Since the social optimum is unlikely to be achieved, it is more fruitful to compare social welfare under different pricing strategies and for the no-piracy benchmark vs. presence of piracy. When network externalities is not important for the pricing decision user shares will be closer to the social optimum when piracy is present relative to the no-piracy benchmark. On the other hand, there will be an efficiency loss from potential quality degradation and from the cost of obtaining illegal copies. It is unclear whether the increased likelihood of network strategic pricing being imposed is welfare enhancing, since it is not generally clear whether network strategic pricing is welfare enhancing.

3.3.1 No network externalities

Relative to the no-piracy benchmark, the profit maximizing price $\tilde{p}_0$, given by equation (17), has two opposing effects on welfare. From proposition 1 the total user share of $W$ is larger relative to the no-piracy benchmark. Hence, the location of the consumer indifferent between $W$ (any type) and $L$, $\bar{m}_0$, is closer to the social optimum: $m_0 \leq \bar{m}_0 < m^*$. However, piracy also entails a welfare loss through a possible degradation of quality, $\alpha$, and the cost of obtaining working illegal copies, $\eta$. In relation to the social surplus in the no-piracy benchmark, welfare with piracy and price $\tilde{p}_0$ is:

$$\tilde{WF}_0 = WF_0 + \int_{\bar{m}_0}^{\tilde{m}_0} (v_W - v_L - \omega)d\omega + (1 - \phi(\tilde{p}_0)) \left[ \int_{-t/2}^{\tilde{m}_0} (-\alpha v_W - \frac{1}{2}\omega)d\omega - \int_0^{\tilde{p}_0 - \alpha v_W} \eta d\eta \right]$$

Whether welfare increases or decreases relative to the no-piracy benchmark is generally ambiguous and depends on a combination of parameters.

3.3.2 Network externalities

The effects on welfare by going from the price $\tilde{p}_0$ to $\tilde{p}_N$ are similar to the no-piracy benchmark. However, since $\tilde{p}_N < \tilde{p}_0$, the piracy rate will be lower, meaning the welfare loss from quality degradation and copying cost is reduced.

Taking pricing incentives into account, it is unclear whether the finding that network strategic pricing is more likely to occur in the presence of piracy is good for welfare. However, if network strategic pricing is imposed due to large product differentiation,
welfare is most likely reduced due to the reduction in the value of users located at large \( \omega \).

From simulations with randomly generated parameter values there is no significant correlation between differences in profits under different pricing strategies, \( \pi(p_N) - \pi(p_0) \), and differences in welfare, \( WF_N - WF_0 \), in both the case with piracy and in the no-piracy benchmark. This suggests that the incentives for engaging in network strategic pricing does not have a clear effect on total welfare. By only inspecting parameter values where network strategic pricing is imposed with piracy, but not without it, welfare increases relative to the no-piracy benchmark in approximately 61% of a total of 1614 simulated model realizations.

4 Analysis and discussion

This analysis aims at identifying whether network strategic pricing is less or more likely to occur with piracy than without. There are two opposing effects present. Since the total network size of \( W \) will be larger and the price lower when piracy is present, the cost from the negative price effect by imposing network strategic pricing is significantly lower due to a smaller price reduction. On the other hand, the positive output effect may also be weaker. By imposing network strategic pricing, the seller of \( W \) may not reap all the benefits from zero demand of \( L \), as a sizable share of former \( L \) users may obtain a pirated version of \( W \) instead. It turns out that the former effect dominates for all parameter values as network strategic pricing usually involves a price so low that the piracy rate will be low, or even zero.

Although the model developed is intuitively simple, the result depends on combinations of parameters. The parameters \( z \) and \( \alpha \) are parameters that are unique to the piracy model, and the values of these two parameters will obviously be important for the results.

The unconditional expected cost-of-copying is \( z/2 \), and intuitively, as \( z \) goes towards infinity, the piracy rate will be zero and we will be back to the no-piracy benchmark. Hence, piracy will only have a substantial effect on the market when \( z \) is not too large, ensuring that there will exist a non-negligible share of new users able to obtain an illegal copy at a low cost.

Similarly, the effect from piracy will be small for large \( \alpha \). In fact, the piracy rate may be zero for quite high prices if illegal copies have much lower quality than original
versions. Hence, a combination of little quality degradation and a large share of users able to copy at a low cost will make piracy a viable option and thus significantly affect pricing incentives for the commercial software developer in competition against both piracy and open source software.

For the numerical analysis, I will assume throughout that the network independent quality of the two products is equal: $v_W = v_L = v$. Although there in reality may be large quality differences in favor of one type, there is in general little evidence that one software type is more superior than the other.\footnote{Lerner and Schankerman (2010), p.5 write: “Based on the studies completed to date, it is hard to draw unambiguous conclusions as to the superiority of open source versus proprietary software.”} In addition, it is assumed that $\lambda = 1$, i.e. that open source software is interpreted as an entrant in a market. Although network strategic pricing may be feasible for a non-negligible positive share of old users, it is unlikely to be an optimal strategy unless the critical mass $N$ is very large. The size of the new and old user population are set to be of equal size: $\sigma = 1/2$.

Figure 6 compares profits under the two pricing strategies in the cases with and without piracy for different values of $N$. The solid lines represent profits under network strategic pricing, while profits under the price $p_0$ where demand for $L$ is above critical mass among new users is represented by the dashed lines. As the figure clearly illustrates, network strategic pricing is optimal under a larger set of values of $N$ when piracy is
present. If the critical mass is interpreted as the strength of network externalities, this means that network strategic pricing will be optimal for weaker degrees of network externalities in the presence of piracy relative to the benchmark case.

Note that the piracy rate is zero under network strategic pricing for all values of $N$ in Figure 6. This means that it is a profit-maximizing strategy to set a price below the no-piracy limit price. The reason for this is that, even at the limit price $p = \alpha v W$, we have $L_0(p) > N$, and profits can be increased further by setting some $p_N < \alpha v W$ which pulls demand for $L$ below the critical mass.

If the share of new users willing to obtain pirated copies falls, i.e., an increase in $z$, or the quality of pirated products is decreased, i.e. $\alpha$ increases, then the critical value of $N$ for network strategic pricing to be optimal will rise, and this critical value will for large values of $z$ and $\alpha$ be equal to that of the no-piracy benchmark.\(^{23}\) That is, the two profit lines in the bottom panel of Figure 6 will never cross at a larger $N$ than what is the case for the no-piracy benchmark in the top panel. Hence, for any set of parameters it cannot be the case that network strategic pricing is optimal in the absence of piracy, while not being optimal when piracy is present.

**Conjecture.** Consider parameter values such that total network sizes exceed the critical mass whenever network strategic pricing is not imposed: $W_0(p_0)$, $W_0(p_0^{\text{piracy}})$, $L_0(p_0)$, $L_0(p_0^{\text{piracy}}) > N$. Then there does not exist a set of parameters such that the network strategic pricing is optimal in the absence of piracy while not being optimal in the presence of piracy: $\pi(p_N) > \pi(p_0)$ and $\tilde{\pi}(\tilde{p}_N) < \tilde{\pi}(\tilde{p}_0)$ is not feasible.

The above conjecture is not formally proven, but is supported by numerical simulations. By repeated draws of random parameter values from large variance uniform distributions no observations found the inequalities $\pi(p_N) > \pi(p_0)$ and $\tilde{\pi}(\tilde{p}_N) < \tilde{\pi}(\tilde{p}_0)$ to hold for the same values of parameters. On the other hand, for a non-negligible share of the draws the opposite was the case, namely that network strategic pricing was not optimal in the absence of piracy, but optimal in the presence of piracy. From Figure 7, a scatter plot of 1297 simulated realizations of the model, it should be noticed that there is not one single observation in the fourth quadrant representing the two inequalities $\pi(p_N) > \pi(p_0)$ and $\tilde{\pi}(\tilde{p}_N) < \tilde{\pi}(\tilde{p}_0)$. Details are provided in the appendix.

The intuition of this result is quite clear. Except for a very limited set of parameters $p_0$ will exceed $\tilde{p}_0$,\(^{24}\) and from Proposition 5 we know that $\tilde{p}_N \geq p_N$. Hence, the price

\(^{23}\)Notice that $\lim_{z \to \infty} \phi(p^*) = 1$ and $\phi(p^*) = 1$ if $p^* < \alpha v W$ for $p^* \in \{p_N, p_0\}$.

\(^{24}\)By running simulations restricting parameters such that the inequality $\tilde{p}_0 > p_0$ holds, the hypothesis from the conjecture was still valid.
reduction by imposing network strategic pricing, $\tilde{p}_0 - \tilde{p}_N$, is reduced, which increases the likelihood of $\tilde{p}_N$ being the profit maximizing price. In addition, other than making the open source competitor worthless, the price $p_N$ will also have a positive effect on sales through a lower piracy rate, which further incentivizes the commercial software developer to impose network strategic pricing.

The importance of the size of the degree of differentiation between $W$ and $L$, measured by the parameter $t$, was discussed in section 3. Figure 8 illustrates how the size of $t$ matters numerically. As $t$ becomes large, a larger share of the new users will value $W$ (and $L$) highly, meaning that the commercial developer can set a high price of the commercial software product when piracy is not present. Hence, the cost of imposing network strategic pricing will be high, as the price differential between $p_0$ and $p_N$ will be large. Therefore, in the absence of piracy, setting the price at $p_0$ is likely to be the profit maximizing strategy for large $t$.

However, if an illegal copy of $W$ is available, the commercial developer is limited in the price-setting. For a given $z$, there is a limit to how high the price can be set, without the firm losing all the customers to piracy. For the numerical example in the bottom panel of Figure 8, $p_0$ ceases to increase as $t$ becomes large. On the other hand, the network strategic price $p_N$ grows convexly in $t$ for $p_N < p_0$, meaning that the cost of imposing network strategic pricing will be very low due to the small price differential between the two pricing strategies. Hence, $p_0$ is likely to be the optimal price for large $t$ (and sufficiently large $N$) in the presence of piracy.
Figure 8: Network strategic pricing not optimal for large degree of differentiation in no-piracy benchmark, but optimal in the presence of piracy
Parameters: $\sigma = 0.5, v = \lambda = z = 1, \alpha = 0.05, N = 0.29.$

5 Empirical evidence

The model predicts that piracy will lead to lower usage of open source software even when network externalities are not present. This is for two reasons: First, when software products are substitutes, the availability of free (or very cheap) copies of commercial software should reduce demand for open source software, as some individuals may choose piracy over open source software. Second, when commercial software developers compete against illegal copies of their own products, they are likely to reduce the price of the legal versions. Lower prices of commercial software will then increase demand for legal copies of commercial software partly at the expense of open source software.

In order to test the effect from piracy on open source software usage empirically, I have obtained data from 104 countries on desktop Linux operating system (OS) user shares in 2011 from the web analytics company Net Applications, as well as estimated piracy rates from the 2011 BSA Global Piracy Study. In addition, GDP data from IMF, and indexes on institutions and infrastructure from the 2012 Global Information Technology Report by the World Economic Forum are used as controls. Descriptive statistics are shown in table 2.

The average national Linux user share in the data set is approximately 1%, but with quite large international variations. While Linux user shares exceed 3% in Finland,
France and Czech Republic, most Middle Eastern countries have Linux users shares below 0.2%. Estimated software piracy rates correlate strongly with per capita income. The United States, Luxembourg and Japan all have estimated piracy rates at around 20%, while in e.g. the Caucasian countries around 90% of all software used is estimated to be illegal copies.

I estimate the following empirical model:

$$\text{Linux usershare}_i = \alpha + \beta \text{piracy rate}_i + \mathbf{x}_i \gamma + u_i$$  \hspace{1cm} (21)

Linux user shares and piracy rates are both measured in logs. Hence $\beta$, the parameter of interest, is interpreted as the elasticity of Linux user share with respect to piracy rate. $\mathbf{x}_i$ is a vector of control variables such as GDP per capita, regional dummies, and internet bandwidth. Internet bandwidth is included as a control since some developing countries suffer from extremely slow internet connections. The primary source of obtaining the Linux installation image is downloading online. Hence obtaining Linux might be costly relative to obtaining an illegal physical copy of Windows. Lack of a functioning internet connection might also be an informational barrier to learning of the existence of software products.

Results from OLS regressions are presented in Table 3. The estimates for $\beta$ are negative and significant, which supports the hypothesis that software piracy has a negative impact on open source usage. The OLS estimates predict that a 1% increase in the piracy rate in a given country on average will result in a reduction in Linux user shares in the area of 0.5-0.6%, depending on the model specification.\(^{25}\)

All the controls have the expected signs. The negative effect from per capita income on Linux usage reflects that the price of Windows relative to income is higher

\(^{25}\)Note that both piracy rates and Linux user shares both are measured in percentages. As an example, for a country with a piracy rate of 50% and a Linux share of 1%, an increase in the piracy rate by 1%, i.e. 0.5 percentage points, is expected to have the Linux user share reduced by 0.005 percentage points assuming $E(\beta) = 0.5$
Table 3: OLS regressions. Dependent variable: ln(Linux share)

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ln(piracy rate)</td>
<td>-0.620***</td>
<td>-0.650***</td>
<td>-0.481**</td>
<td>-0.627**</td>
</tr>
<tr>
<td></td>
<td>(0.202)</td>
<td>(0.198)</td>
<td>(0.235)</td>
<td>(0.261)</td>
</tr>
<tr>
<td>ln(GDP/cap (PPP))</td>
<td>-0.344***</td>
<td>-0.296***</td>
<td>-0.420***</td>
<td>-0.150</td>
</tr>
<tr>
<td></td>
<td>(0.111)</td>
<td>(0.0975)</td>
<td>(0.135)</td>
<td>(0.125)</td>
</tr>
<tr>
<td>ln(internet bandwidth)</td>
<td>0.0587</td>
<td>0.282***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0555)</td>
<td>(0.0717)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asia Pacific</td>
<td>-0.791***</td>
<td>-0.820***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.141)</td>
<td>(0.148)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mid. East and Africa</td>
<td>-1.055***</td>
<td>-1.127***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.176)</td>
<td>(0.159)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-1.762*</td>
<td>-2.053**</td>
<td>-1.954*</td>
<td>-3.845***</td>
</tr>
<tr>
<td></td>
<td>(0.921)</td>
<td>(0.850)</td>
<td>(1.111)</td>
<td>(1.053)</td>
</tr>
<tr>
<td>Observations</td>
<td>104</td>
<td>104</td>
<td>104</td>
<td>104</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>0.412</td>
<td>0.414</td>
<td>0.154</td>
<td>0.036</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

In developing economies, and poorer individuals are therefore more inclined to demand free-of-charge substitutes of expensive software. As expected, higher internet speed will on average lead to higher Linux usage. However, the estimate is only significant when regional dummies are not controlled for.

It cannot be ruled out that the demand for pirated Windows copies depends on usage of Linux, meaning that the piracy rate might be correlated with the error term in the empirical model. If more people use open source software, this might also lead less people to obtain pirated software. Using an index of intellectual property (IP) protection as an instrument for the piracy rate, I find results very similar to the OLS estimates in table 3. Since usage of pirated software is an illegal activity, it can be argued that the degree of IP protection will have a negative impact on the extent of piracy. Likewise, since there hardly are any (enforceable) legal restrictions on the use of Linux,\(^{26}\) it is unlikely that the degree of IP protection directly affects demand for Linux.\(^{27}\) Nevertheless, the

---

\(^{26}\)As described in the introduction, open source licenses may e.g. have restrictions on license type of modified open source products.

\(^{27}\)An objection to this argument may be that countries with high IP protection also are countries likely to have a large ICT sector. Since Linux is more widely used in the ICT sector, IP protection may no be a valid instrument if it correlates with Linux usage.
Hausman (1978) test supports the null hypothesis that the OLS estimates are consistent. The analysis using IP protection as instrumental variable is available in the appendix.

These results are consistent with the findings of the brief empirical analysis conducted by Casadesus-Masanell and Ghemawat (2006). By using data on shipments of server operating systems in 45 countries they find that a country’s piracy rate has a negative effect on the Linux/Windows shipment ratio (significant at the 10% level) controlled for per capita income. While they estimate the expected negative effect, their data is problematic. Considering that any Linux OS is free of charge and can easily be downloaded online, shipments of Linux OSs is likely to be an inaccurate proxy of actual Linux usage. Also, using a ratio between two OSs as a dependent variable may be sensitive to variations of third types of software products, such as Apple’s OS-X in the case of operating systems. By using the same empirical strategy as Casadesus-Masanell and Ghemawat (2006) with the data from the above analysis, I estimate a negative coefficient significant at the 5% level, consistent with their findings.

For the sake of comparison I conducted a similar analysis on the relationship between software piracy and usage of Apple’s OS-X across countries. Since OS-X is bundled with hardware (Macbook and iMac), high-quality pirated copies of Apple computers are hard to come by at a low price. Therefore, Apple is not plagued by piracy of their operating systems to the same degree as software developers like Microsoft who primarily sell software on third-party hardware. For this reason, one would expect that a high software piracy rate in a given country would have a negative impact on usage of OS-X. The OLS regression estimates that a 1% increase in the piracy rate would on average lead to a decrease by around 0.95% in the OS-X user share (regression table not included). The effect from piracy on OS-X being greater than what was estimated for Linux may be explained by OS-X being more mature than Linux with regards to installed base and therefore being endowed with a larger portfolio of compatible applications. With regards to the interpretation of network externalities in the theoretical model in this paper, few individuals consider Linux to be above critical mass, and will therefore not wish to obtain Linux for free, even if obtaining illegal copies of Windows is costly. However, the opportunity cost of purchasing an Apple computer is significantly lower when pirating Windows is out of the question.

Original hardware products are often built by expensive components, and illegal replicas tend to be made from inferior components to save costs. Moreover, proprietary components such as the motherboard are close to impossible to perfectly replicate by outsiders.
This informal comparison on the effect from piracy on usage of Linux and OS-X, respectively, may suggest that one could expect larger coefficients on piracy’s impact on Linux usage had Linux been more widely used initially. Ideally, a similar cross-country analysis should be conducted of a more widely used open source software product. The office suites LibreOffice and OpenOffice are likely to be good candidates as these products compete with a dominant Microsoft Office, but do still maintain a non-negligible market share.\textsuperscript{29}

6 Conclusion

This paper has analyzed optimal pricing strategies when a proprietary software developer faces competition from both piracy and free-of-charge open source software in a market with network externalities. Two pricing strategies for the commercial software provider are identified: (i) ”standard” pricing where the network size of the OSS product is not strategically accounted for in the pricing decision, and (ii) ”network strategic pricing” where the price of the commercial software product is deliberately set with the intention of a critical mass of users not being attained for the OSS type. The result of imposing network strategic pricing is a market where no one will choose to download open source software, and commercial software products will therefore dominate the market.

The main finding from the theoretical model is that the conditions for network strategic pricing are more easily satisfied in the presence of piracy relative to the no-piracy benchmark. The explanation is twofold: First, piracy of commercial software will add to the network of the commercial software type at the expense of open source software. When the installed base of OSS users is initially smaller, the cost of imposing network strategic pricing is reduced, and therefore it is more likely to be optimal. Second, competition from illegal piracy will force the developer of commercial software to cut the price, which will increase the number of purchasing customers, partly at the expense of the installed base of OSS. Through this channel, the cost of introducing network strategic pricing is reduced further.

Network strategic pricing is more likely to be optimal when network externalities are strong, independently of the existence of illegal copies of commercial software. A low cost of copying, translating into a price-sensitive piracy rate, combined with quality of

pirated products close to that of the original are factors that increase the likelihood of network strategic pricing being imposed. In addition, a large degree of differentiation between open source and commercial software makes network strategic pricing unlikely when piracy is absent, but very likely in the presence of piracy. In a world without piracy, product differentiation creates market power translating into high prices, while commercial software providers lose market power if consumers are likely to choose piracy over purchasing if prices are too high. Hence, with piracy the price differential between the two pricing strategies is small when product differentiation is large, and the commercial software provider can force open source software out of the market with a relative small price reduction.

Since software goods consist of information only, unauthorized replication usually leads to little loss in quality. In addition, price is arguably the most important factor explaining the extent of software piracy. Although piracy rates vary substantially between countries, the piracy rate is high for commercial software products in general, which suggests that the average cost of copying is fairly low for most software products. Hence, network strategic pricing is likely to be an optimal pricing strategy for commercial software products benefiting from a large installed base relative to OSS substitutes.

The theoretical model might provide an explanation for Microsoft’s pricing of products like Windows and Office. Despite their dominant market position and a huge army of faithful users, their products are not priced nearly as high as some software products provided by other dominant developers of proprietary software (e.g. Adobe). Microsoft products are associated with strong network externalities, easily accessible illegal copies, and large installed bases. Substitutes like OpenOffice and various types of Linux operating systems are generally regarded as being of high quality, but are also little used. It can therefore be argued that it is a strategic pricing decision by Microsoft not to price their products too high in order to prevent open source competition to increasing its installed base, thus preventing network benefits being obtained.

The model might also give an additional rationale for the existence of student discounts on software. Being both experimental and price-sensitive, students might be more likely to try out free alternatives to proprietary software products. Hence, proprietary

---

software developers might be incentivized to lower the price in order to avoid open source competition in attaining a critical mass of users. Arguably, students and young people in general have lower costs, both technically and ethically, in obtaining pirated products. Hence, the existence of high-quality illegal copies will incentivize proprietary developers further to impose network strategic prices to students.

The results from the theoretical model in this paper suggest that piracy is hurtful to both proprietary and open source software developers. While piracy increases usage of commercial software, illicit copying is likely to affect profits negatively through lower prices and possibly reduced sales. Piracy reduces usage of open source software both through the direct effect of individuals leaving open source software products in favor of free-of-charge illegal copies, as well as the indirect effect of lower prices on commercial alternatives. In addition, the presence of piracy may more easily incentivize proprietary software developers to lower prices significantly in order to avoid free-of-charge competitors to generate the network externalities needed to create demand. Although not being directly affected by software piracy, open source enthusiasts have good reasons to advocate for increased copyright protection.

References


A Appendix

A.1 Proofs

Proof of proposition 2: By inserting \( p_0 \) from equation (3) into the LHS of equation (18), i.e. the marginal profit function \( \tilde{\pi}'(p) \), the result is proven if there exist parameter values such that \( \tilde{\pi}'(p_0) > 0 \).

Note that if \( \phi(p_0) = 1 \), then \( p_0 = \tilde{p}_0 \) since this implies a zero piracy rate. If \( \phi(p_0) = 0 \), then \( p_0 > \tilde{p}_0 \) since this price will yield a piracy rate of 100% and zero profits, thus it must be optimal to set a price lower than \( p_0 \).

Hence, we must restrict the parameters such that \( 0 < \phi(p_0) < 1 \), which will be satisfied if we set \( p_0 = \alpha v_W + h z \), where \( h \in (0, 1) \). By inserting (3) for \( p_0 \) we restrict \( t \) to be of the form:

\[
t = 4(\alpha v_W - \frac{v_W - v_L}{2} + h z)
\]

It can be shown that \( \tilde{\pi}'(p_0) \) now can be written as:

\[
\tilde{\pi}'(p_0) = -\frac{\sigma}{2 t z} \left( 2 \alpha^2 v_W^2 + 3 \alpha v_W h^2 z + 2 \alpha v_W h z + 4 h^3 z^2 - h^2 z^2 \right).
\]

Hence, \( \tilde{\pi}'(p_0) > 0 \) if

\[
2 \alpha^2 v_W^2 + 3 \alpha v_W h^2 z + 2 \alpha v_W h z + 4 h^3 z^2 - h^2 z^2 < 0,
\]

then there will exist parameter values where \( \tilde{p}_0 > p_0 \) if the set \( (0, -2 \alpha^2 v_W^2 - \alpha v_W h z (3h + 2) + h^2 z^2) \) is non-empty. In order to show that a non-empty set may exist within the model’s assumptions, we can set \( \alpha = 0 \). Then \( \tilde{\pi}'(p_0) > 0 \) if \( 0 < h < \frac{1}{4} \).

\[\square\]

Proof of proposition 5 (alternative method): The proof consist of two parts. The second part derives from the following lemma:

(i) Lemma A1: There exists a unique pair \((\tilde{L}_0^{\text{new}}(p), p)\) s.t. \( \tilde{L}_0^{\text{new}}(p, t_i) = \tilde{L}_0^{\text{new}}(p, t_j) \) for any pair \( t_i, t_j, i \neq j \)

Proof of lemma We have that

\[
\frac{\tilde{L}_0^{\text{new}}(p)}{\sigma} = \frac{1}{2} \left[ \frac{\phi(p)(v_W - p) + (1 - \phi(p))(1 - \alpha)v_W - v_L}{t} \right]
\]

where

\[
\phi(p) = \min \{1 - \frac{p - \alpha v_W}{z}, 1\}.
\]

Assuming \( p > \alpha v_W \), we find the point where two demand curves cross for two arbitrary values of \( t \) (for \( p < \alpha v_W \) calculations are straight forward):

\[
\tilde{L}_0^{\text{new}}(p, t_i) = \tilde{L}_0^{\text{new}}(p, t_j), t_i \neq t_j
\]
Solving for \( p \) and adding the restriction that \( \frac{\partial L_{0}^{\text{new}}(p)}{\partial p} > 0 \), i.e. that \( \phi(p) > 0 \), we find the solution:
\[
\bar{p} = \alpha v_W + z - \sqrt{z \left[ 2(v_L - (1 - \alpha)v_W) + 2 \right]}
\]
which is independent of \( t \).

(ii) We have
\[
\frac{\partial \tilde{L}^{\text{new}}_{0}(p)}{\partial p} = \frac{\sigma}{t} \phi(p)
\]
which is positive when \( \phi(p) > 0 \). Hence, the larger \( t \) is, the steeper is the inverse demand curve. From (i) we know that two demand curves with different values of \( t \) will cross at the point \( (L_{0}^{\text{new}}(\bar{p}), \bar{p}) \). Hence, for \( t_1 > t_2 \), we must have that \( L_{0}^{\text{new}}(p, t_1) < L_{0}^{\text{new}}(p, t_2) \) for \( p > \bar{p} \). From this it follows that \( \tilde{p}_N(t_1) > \tilde{p}_N(t_2) \) for \( N - L_{\text{old}}^{L} > \tilde{L}_{0}^{\text{new}}(\tilde{p}) \). See figure 9 for a graphical argument.

![Figure 9: \( \tilde{p}_N \) increasing in \( t \) for \( N - L_{\text{old}}^{L} > \tilde{L}_{0}^{\text{new}}(\tilde{p}) \), \( t_1 > t_2 \).](image)

\( \Box \)

**Evidence of conjecture 1**

Conjecture 1 is not formally proven, but simulations from randomly drawn parameter values give strong statistical evidence in favor of the conjecture.

I draw parameter values from the following uniform distributions: \( \sigma \in [0.1, 0.7] \), \( v_W \in [0.1, 10] \), \( v_L \in [0.1, 10] \), \( t \in [0.01, 10] \), \( \lambda \in [0.5, 1] \), \( N \in [0.2, 0.35] \), \( \alpha \in [0, 0.5] \), \( z \in [0, 10] \).
Figure 10: Scatter plot of 1297 simulations. \( \pi(p_N) > \pi(p_0) \) and \( \tilde{\pi}(\tilde{p}_N) < \tilde{\pi}(\tilde{p}_0) \) does not occur for the same parameter values.

The model was simulated 20 000 times, and the restrictions \( W_0(p_0), \tilde{W}_0(\tilde{p}_0), L_0(p_0), \tilde{L}_0(\tilde{p}_0) > N \) as well as \( \phi(\tilde{p}_0) < 1 \) were true 1297 times of these draws.\(^{31}\)

Figure 10 plots the 1297 random draws with \( \pi(p_N) - \pi(p_0) \) on the horizontal axis and \( \tilde{\pi}(\tilde{p}_N) - \tilde{\pi}(\tilde{p}_0) \) on the vertical axis. From the figure it is clear that there is not one single observation where \( \pi(p_N) > \pi(p_0) \) and \( \tilde{\pi}(\tilde{p}_N) < \tilde{\pi}(\tilde{p}_0) \) is true for same parameter values, while for a non-trivial share (exactly 178 observations) of the simulated data \( \pi(p_N) < \pi(p_0) \) and \( \tilde{\pi}(\tilde{p}_N) > \tilde{\pi}(\tilde{p}_0) \). I.e. the commercial software provider imposes network strategic pricing as a response to piracy being present.

### A.2 Instrumental variable estimation

A potential problem with the estimation of the empirical model in section 5 is that the piracy rate might be correlated with the error term. As argued, demand for Linux operating systems may be negatively affected by the piracy rate since individuals may obtain pirated products instead of downloading Linux for free if pirated versions of Windows are easily accessible (or indirectly through lower prices of Windows due to competition from pirated copies). However, the causality may be the reverse: Because individuals download Linux for free, they choose not to obtain illegal pirated copies of Windows.

In order to correct for this possible endogeneity problem, I use a measure of intellectual property (IP) protection as a instrument for the piracy rate in a given country

\[^{31}\text{The restriction of } \phi(\tilde{p}_0) < 1 \text{ was included to ensure that the piracy model and no-piracy benchmark gave different results.}\]
(obtained from the 2012 Global Information Technology Report by the World Economic Forum). The measure of IP protection is a score between 1 and 7, where 7 is the highest level of IP protection possible.

With regards to the validity of the IV, it is likely that IP protection has a strong negative effect on piracy rate in a given country, since piracy is a clear violation on IP laws. On the other hand, it is unlikely that the degree of IP protection has any direct effect on Linux usage as hardly any IP laws apply to usage of open source software. However, IP protection is likely to be correlated with the importance of the ICT sector. Since Linux is more widely used in the ICT sector than other sectors, the validity of IP protection as IV may correlate with Linux usage. Including internet bandwidth as a control may mitigate this potential problem somewhat since good internet infrastructure is of great importance for a functional ICT sector.

Table 4 shows the 2SLS regression results. The estimates turn out to be close to those in the OLS regressions from section 5. When regional dummies are included, the effects from piracy on Linux usage are estimated to be larger using IV relative to OLS, and are not significant when regional dummies are omitted. However, the standard errors are also larger, and the larger coefficient might be due to small sample bias of 2SLS. The first stage is very strong with F-values varying between 94 and 113. Nevertheless, the null hypothesis of consistent OLS estimates is not rejected for any of the model specifications, which suggest that the OLS is the preferred estimation method.
Table 4: 2SLS regression. Instrument for ln(piracy rate): Intellectual property protection (index: 1-7). Dependent variable: ln(Linux share)

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ln(piracy rate)</td>
<td>-0.998***</td>
<td>-1.038***</td>
<td>-0.138</td>
<td>-0.231</td>
</tr>
<tr>
<td></td>
<td>(0.350)</td>
<td>(0.340)</td>
<td>(0.426)</td>
<td>(0.443)</td>
</tr>
<tr>
<td>ln(GDP/cap (PPP))</td>
<td>-0.447***</td>
<td>-0.422***</td>
<td>-0.325*</td>
<td>-0.0217</td>
</tr>
<tr>
<td></td>
<td>(0.133)</td>
<td>(0.129)</td>
<td>(0.168)</td>
<td>(0.161)</td>
</tr>
<tr>
<td>ln(internet bandwidth)</td>
<td>0.0384</td>
<td>0.296***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0565)</td>
<td>(0.0743)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asia Pacific</td>
<td>-0.819***</td>
<td>-0.838***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.149)</td>
<td>(0.158)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mid. East and Africa</td>
<td>-1.077***</td>
<td>-1.124***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.175)</td>
<td>(0.153)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-0.961</td>
<td>-1.102</td>
<td>-2.681**</td>
<td>-4.812***</td>
</tr>
<tr>
<td></td>
<td>(1.074)</td>
<td>(1.051)</td>
<td>(1.342)</td>
<td>(1.296)</td>
</tr>
<tr>
<td>Observations</td>
<td>104</td>
<td>104</td>
<td>104</td>
<td>104</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>0.394</td>
<td>0.394</td>
<td>0.139</td>
<td>0.016</td>
</tr>
<tr>
<td>First stage IV F-value</td>
<td>97.50</td>
<td>112.99</td>
<td>94.45</td>
<td>94.63</td>
</tr>
<tr>
<td>Hausman test p-value</td>
<td>0.20</td>
<td>0.17</td>
<td>0.35</td>
<td>0.31</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$