View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by AIS Electronic Library (AISeL)

Association for Information Systems
AIS Electronic Library (AISeL)

ICIS 2001 Proceedings

International Conference on Information Systems (ICIS)

December 2001

Opening the Code: How Open Is Optimal?

Geoffrey Parker
Tulane University

Marshall van Alstyne
University of Michigan

Follow this and additional works at: http://aisel.aisnet.org/icis2001

Recommended Citation

Parker, Geoffrey and van Alstyne, Marshall, "Opening the Code: How Open Is Optimal?" (2001). *ICIS 2001 Proceedings*. 65.
http://aisel.aisnet.org/icis2001/65

This material is brought to you by the International Conference on Information Systems (ICIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in ICIS 2001 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

OPENING THE CODE: HOW OPEN IS OPTIMAL?

Geoffrey Parker
Freeman School of Business
Tulane University
gparker@tulane.edu

Marshall Van Alstyne
School of Information
University of Michigan
mvanalst@umich.edu

Abstract

Recent developments have challenged one prevailing interpretation of the idea that proprietary systems, enshrined in copyright, create the greatest value. The challenge appears at one level among economic strategists who assert that the greatest value in information goods is not created by the strongest and most restrictive intellectual property protection and in another form by the proponents of Open Source Software who argue for value created by peer review and openly modifiable, shared code. We articulate a balance of incentives and openness to promote both the creation of new products and the network externality benefits from open access. We consider the welfare of consumers and producers to show that environmental parameters such as the size of the market, the network effects, and the locus of innovation can affect the optimal choice of time to release and degree of openness.

Keywords: Open source software, GNU Public license, network externality, software development.

INTRODUCTION

Economists have long argued that incentives matter. Intellectual property law embodies this principle, granting temporary monopolies to authors and inventors as a stimulus to innovation. Recent developments, however, have challenged one prevailing interpretation of this idea that proprietary systems, enshrined in copyright, create the greatest value. The challenge appears at one level among economic strategists who assert that the greatest value in information goods is not created by the strongest and most restrictive intellectual property protection (Shapiro and Varian 1999). It appears in another form among proponents of the Free Software Foundation, who argue for openness as a right, and the Open Source Software movement, who argue for value created by peer review and network benefits of openly modifiable, sharable code.

This paper articulates a balance of incentives and openness that promotes both the genesis of feature rich new software offset by widespread distribution and network externality benefits from open access. Using analysis and simulation, we find that each system has merits and that conditions exist when one or the other may dominate. Considering the welfare of consumers and producers, who may at times represent one and the same market, we also show that developer interests such as code reuse and environmental parameters such as the size of the market can affect the optimal choice of when to release proprietary restrictions and the degree of openness.

RESEARCH QUESTIONS

The specific research questions that we address in this work are as follows:

- What fraction of a code base should a developer release immediately? When, if ever, should all code be released?

- When, if ever, should third party enhancements be released back to the open code base?

- What impact does the size of the developer community have on the amount of code released and the timing of release?
- How is consumer surplus and producer surplus (i.e. , social welfare) affected by different release policies?

The question of how much open code to release is a decision that authors can make when designing a successful information good. At time zero, authors can choose to release some fraction of their code base in order to foster immediate adoption and allow innovation on their platform. They can also set a future date at which all of a code base is released. If the developer can create complementarity between the code base and a good in which the developer has pricing power, then there is an incentive to release a substantial fraction of the code base immediately. However, code that is immediately released does not contribute directly to the author's profits. Moreover, promoting subsequent innovation via third party enhancements faces the same trade-off: Tension arises between promoting network growth by earlier release and promoting innovation by later release.

THEORETICAL FOUNDATIONS

We analyze the trade-offs based on a modeling innovation that accounts for *inter*-network externalities, grounded in a two-sided market. This phenomenon has only recently been addressed in a pair of papers Parker and Van Alstyne (2000) and Rochet and Tirole (2001). In a twist on the idea of a network externality—a demand economy of scale evident, for example, in telephones and fax machines—this considers a demand economy of scale that crosses markets as distinct from one that stays within the same market. A two-sided market is evident, for example, to E-Bay as coupled buyers and suppliers, to VISA as member banks and merchants, and to web portals as advertisers and content providers. The presence (or absence) of each side makes the other more (or less) valuable to an organization that sells to both halves at once.

In the present context, software developers *D* create enhancements that add value for end-users or consumers *C*. The more that developers are motivated to create enhancements, the greater the benefits to *C*. Similarly, the larger the *C* market, the more attractive it is for developers to code for that market, contingent on their access to source-code bases. This is consistent with the size of a problem scope influencing the prestige associated with solving it (Raymond 1998). The computer games market represents one example where the presence of a large number of free game levels makes it attractive to consume particular games; and the presence of a larger gaming user community makes it attractive to develop for a particular game. Similarly, for operating systems, the larger the installed base of users, the more attractive is developing applications for that OS; and the more applications, the more attractive it is to consume that OS.

We begin with a simple model. Let parameters Q and V represent bounds on market size and value in the absence of externalities. Further, allow subscripts c and d to index consumer and developer markets respectively. An externality term e_{cd} determines how much of an impact consumption in the developer market has on the consumer market and similarly for e_{cd}. This yields the pair of simultaneous equations:

1. $q_c = Q_c + e_{dc} q_d - \frac{Q_c}{V} p$

2. $q_d = Q_d + e_{cd} q_d - \frac{Q_d}{V} p$

Eliminating q_i from the right-hand-side in each equation leads to independent formulae. These we can use to calculate sales, profits, and consumer surplus.

$$q_c = \left(\frac{1}{1 - e_{cd} e_{dc}}\right)\left(Q_c\left(1 - \frac{p_c}{V}\right) + e_{dc} Q_d \left(1 - \frac{p_d}{V}\right)\right)$$

$$q_d = \left(\frac{1}{1 - e_{cd} e_{dc}}\right)\left(Q_d\left(1 - \frac{p_d}{V}\right) + e_{cd} Q_c \left(1 - \frac{p_c}{V}\right)\right)$$

$$CS_c = \int_0^{q^*} \frac{V(Q_c - q_c + e_{dc} q_d)}{Q_c} dq_c$$

Once we impose concavity restrictions on the profit function $\pi = p_c q_c + p_d q_d$, optimal prices and profits can be found analytically. For quantities, the solution to this system of equations includes a network externality term or "market multiplier" effect $M = \frac{1}{1 - e_{cd} e_{dc}}$ (natural first order concavity restrictions also impose $0 \le e_{cd} e_{dc} < 1$). As the size of the inter-network externality term shrinks, the market behaves more like a normal tangible good without network effects. On the other hand, as $e_{dc} e_{cd} \rightarrow 1$, M becomes arbitrarily large and network effects become a serious consideration in determining an optimal license.

We proceed as follows. Let σ represent a sharing parameter that governs the fraction of code released under open source, which is no longer purchased but is freely adopted by consumers and developers. We also introduce a proprietary time period t that generates revenues for third party developers on any enhancements they create. Before t expires, enhancements are profitable but consumer adoption is also limited via positive prices. Upon expiration of t, developers contribute derivative works back to the open code base from which they drew their original sources. Finally, we endogenize total value V as the sum of value created by the original author V_a and by subsequent developers $V_{d1}, V_{d2},...V_{dn}$. Developer value is increased by the innovation incentive based on how long they can charge for their enhancements. We then explore optimal choices for t and σ based on different welfare maximizing criteria, including combinations of freedom of access, adoption rates, and revenues.

To see the welfare value of enhancements, consider Figure 1.

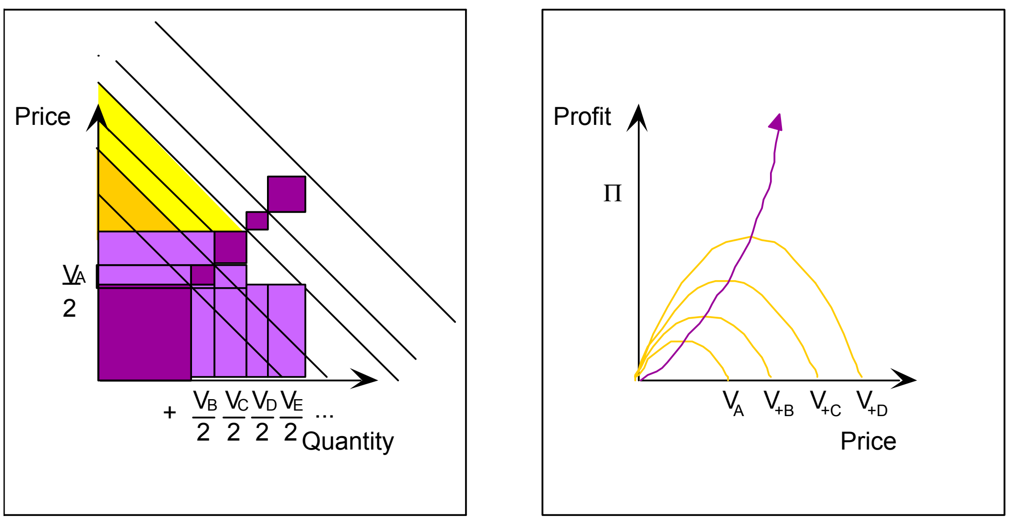

Figure 1. Network Profit Increases Non-linearly with Subsequent Innovation

Given initial value V_a, producer and consumer surplus are depicted in the figure on the left. Enhancements $V_{d1}, V_{d2},...V_{dn}$ create value and potential new sales even if the original author adds no new value directly. The extent to which this occurs is governed by the shared code subsidy governed by σ and the residual profits to developers governed by t.

Not all results are analytically solvable so we use simulation to characterize optimal outcomes under different parameter values. One important contribution of this framework is that it generalizes several different types of contracts. We can then ask which contract optimizes a given welfare criterion under a specific set of parameter values. If we vary the degree of openness $\sigma \in [0,1]$, we can model a range of access to source code that spans, for example, closed proprietary systems, access to application program interfaces (APIs), trial-ware, share-ware, royalty free binaries, open source (BSD), and viral open source (GNU). Similarly, different times to release $t \in [0,1]$ can represent disclosure now, after a brief delay, after a long delay, or on expiration of copyright. Several existing licenses are then captured in Figure 2.

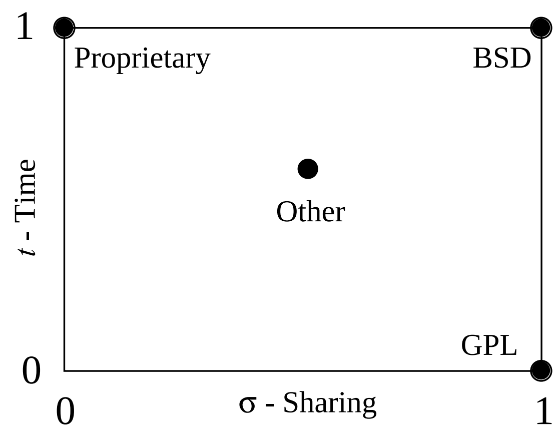

Figure 2. Contracts Mapped by Time-to-Release (0 = Now, 1 = Never) and Sharing (0 = None, 1 = Complete)

Fully open source software, such as that released under the GNU Public License (GPL) releases everything, $\sigma = 1$, and requires subsequent code to be released immediately, $t = 0$. In contrast, the Berkeley Standard Developer's (BSD) license releases everything but then places no restrictions on subsequent disclosure, $t = 1$. A completely proprietary system releases no open software, $\sigma = 0$, and also does not require disclosure, $t = 1$. Ironically, these appear as corner solutions and there are also intermediate licenses, which we propose, such that an optimum occurs at $0 < \sigma < 1$ and $0 < t < 1$.

The parameters σ and t influence critical trade-offs. An increasingly open license, with σ rising from 0 toward 1, has two opposing effects: (1) it promotes adoption as a fraction of the software value is given away for free and (2) it diminishes profits as this fraction of total value no longer generates profits. The first effect increases the size of the network, increasing potential profits, while the second diminishes sales, decreasing potential profits.[1]

The software author incurs a similar set of trade-offs in choosing an optimal t governing the time that subsequent enhancements must be released. As t rises from 0 toward 1, it causes (1) a longer delay until proprietary innovations are returned to the open code base and (2) a rise in the investment incentive for subsequent innovations due to the greater time to collect revenues. The first effect decreases the size of the network, and lowers potential profits, by limiting adoption to paying customers. The second effect increases the value of the total product, raising potential profits, by increasing complementary investments in the overall product.

Equations one and two specify a static model. To capture temporal effects, we use a two period model and a discount factor that allows consumers to choose purchases either in period one, where they enjoy benefits sooner, or in period two, where they enjoy open code at a lower price. The original firm then must factor consumers strategic behavior into determining the optimal time to release the code under open code terms. Forcing early release of enhancements creates consumer surplus at the cost of reducing developer incentives. Numerical simulation is used to explore profit and welfare maximization tradeoffs that also meet consumer equilibrium conditions.

Within this framework, we illustrate which license maximizes a given welfare criterion and then determine the environmental state variables for when each license makes sense. Using a standard welfare criterion, the sum of producer and consumer surplus, the primary licenses are optimal under different criteria. As research in progress, these results are preliminary

1. *Proprietary*: In particular, if an original author can create enhancements on his/her own and the network externality effect is negligible, then there is little benefit from opening the code, so $\sigma = 0$. Furthermore, because the original author

[1]We note that "free" in the sense of freedom is intended by the GPL and not in the sense of price. Importantly, however, the right of anyone to freely redistribute copies of original sources also implies that an author must compete with zero marginal cost copies of his or her own code. Price competition in such a market naturally forces prices toward marginal cost, here zero, so "free" in the sense of price appears reasonably defensible.

has released no source code, the absence of leverage with subsequent developers implies that the proprietary period extends to the duration of copyright and $t = 1$. Original author surplus is maximized. Total welfare is maximized if the author would not have brought the product to market without his or her own incentive period.

2. *BSD*: If the consumer market offers little value relative to the developer market s.t. $Q_c << Q_d$ and there are moderate network effects $M > 0$, then there is more benefit from opening the source code than from keeping it closed and $\sigma = 1$. If innovation is important so that the original author's contribution V_a is small relative to $V_{d1} + V_{d2} \ldots V_{dn}$, then innovators need to be rewarded and $t = 1$. Since the paying market is small but innovators need to be rewarded, a BSD license appears to maximize welfare under these parameter conditions.

3. *GPL*: If the consumer market is either negligible or the same as the developer market, which consumes its own enhancements, and network effects are significant, then GPL appears to maximize welfare.

Using a non-standard welfare criterion, there appears to be another occasion when GPL dominates other licenses. If the goal is to strictly minimize the amount of proprietary code that "prevents one programmer from sharing with another," then setting both $\sigma = 0$ and $t = 0$, opens the original source code and subsequent enhancements. Using this criterion, the GPL is the dominant license. However, if the criterion is modified slightly and defined as the total value of open code available to programmers, then brief periods of proprietary licensing in the short run can create greater welfare in the long run.

Interestingly, it is also possible to define licenses with interior solutions for which $0 < \sigma < 1$ and $0 < t < 1$. Such a license might be termed a flexible copyright, or meta-license, that incorporates features of both proprietary copyright and open source copyleft. These licenses can potentially maximize welfare when externalities are significant, innovation is significant, and users are distinct from developers.

References

Boyle, J. *Shamans, Software and Spleens: Law and the Construction of the Information Society*, Harvard University Press, Cambridge, MA, 1996.

Chen, Y., and Png, I. "Software Pricing and Copyright Enforcement: Private Profit vis-à-vis Social Welfare," in *Proceedings of the Twentieth International Conference on Information Systems*, P. De and J. I. DeGross (eds.), Charlotte, NC, 1999.

Conner, K., and Rumelt, R. "Software Piracy: An Analysis of Protection Strategies," *Management Science* (37:2), 1991, pp. 125-140.

Davis, R., Samuelson, P., Kapor, M., and Reichman, J. "A New View of Intellectual Property and Software," *Communications of the ACM* (39:3), 1996, pp. 21-30.

Galler, B. *Software and Intellectual Property Protection*, Quorum Books, Westport, CT, 1995

Gilbert, R., and Shapiro, C. "Optimal Patent Length and Breadth," *Rand Journal of Economics* (21:1), 1990, pp. 106-112.

Hecker, F. "Setting Up Shop: The Business of Open-Source Software," *IEEE Software*, January/February 1999, pp. 45-51.

Lerner, J., and Tirole, J. "The Simple Economics of Open Source," NBER Working Paper 7600, National Bureau of Economic Research, Cambridge, MA, 2000.

Lessig, L. Code and other Laws of Cyberspace,. Basic Books, New York, 1999.

Merges, R. "The End of Friction? Property Rights and Contract in the 'Newtonian' World of On-Line Commerce," *Berkeley Technology Law Journal* (12), 1997, pg. 115.

Moglen, E. "Anarchism Triumphant: Free Software and the Death of Copyright" *First Monday* (4), 1999.

Parker, G. ,and Van Alstyne, M. "Information Complements, Substitutes, and Strategic Product Design," Social Sciences Research Network Working Paper, 2000 (www.ssrn.com).

Raymond, E. "The Cathedral and the Bazaar," 1996 (available at http://hurkle.thyrsus.com/~esr/writings/).

Raymond, E. "Homesteading the Noosphere," 1998 (available at http://hurkle.thyrsus.com/~esr/writings/).

Raymond, E. "The Magic Cauldron," 2000 (available at http://hurkle.thyrsus.com/~esr/writings/).

Rochet, J. C., and Tirole, J. "Platform Competition in Two-Sided Markets," 2001 (www.idei.asso.fr/Commun/Conferences/CSIO-IDEI/2001/Papiers/Rochet.pdf).

Samuelson, P. "Is Information Property?," *Communications of the ACM* (34:3), 1991, pp. 15-19.

Shapiro, C., and Varian, H. R. *Information Rules: A Strategic Guide to the Information Economy*, Harvard Business School Press, Boston, 1999.