Economics of Open Source: A Dynamic Approach*

Jeongmeen Suh† Murat Yılmaz‡

KIEP Boğaziçi University

March 2009

Abstract

This paper analyzes open innovation projects and their effects on incentives for innovation. We model basic features of the General Public License (GPL), one of the most popular open source licenses and study how firms behave under this license. Under the GPL, there is a trade-off between stimulating innovation and promoting disclosure. By using the open source, a firm can increase its technology level and therefore its probability of innovation success and of achieving a greater profit in that period. However, any innovative findings using open source would be also open source in subsequent periods. This obligation decreases the expected future revenue of the firm. We analyze this trade-off and show that if a firm has the same technology level as the open source, it does not use the source. On the other hand, if a firm has a lower level of production technology than the open source, it is optimal to use the source.

*We are grateful to Barton Lipman for his advice and support. We would also like to thank Andrew Newman, Dilip Mookherjee and Ching-to Albert Ma for their useful feedback.

†KIEP, 108 Yangjaedaero Seocho-gu Seoul, Korea. Phone: +82 2 3460 1114

‡Boğaziçi University, Department of Economics, Bebek, Istanbul, 34342. Email: *muraty@boun.edu.tr*. Phone: +90 212 3597656. Fax: +90 212 287 2453. Web site: *http://www.econ.boun.edu.tr/yilmaz/*

1 Introduction

Open software development involves a major deviation from the private investment model of innovation; open source innovators freely share the proprietary software that they have developed at their private expense. For example, Linux, a computer operating system, is evolving with many independent developers revealing the code to develop and refine it. Its source code is *open* in the sense that anyone has free access to it. The success of open source software raises many questions about innovation policies with non-traditional property rights.

In this paper, we study a major feature of open innovation and its effects on incentives for innovation. Although the source code of open source software is freely available, open source programs are distributed under very precise licensing agreements. The GNU General Public License (GPL) is one of the most common licenses and we model its key characteristic; specifically while every user has the freedom to use and modify programs subject to the GPL, such modifications must be distributed under the terms of the license itself if they are to be distributed at all.

The GPL, however, does not preclude the commercial exploitation of the software, at any stage. That is, the program users have to maintain the free access to the source, but they do not need to share any profit they make. There are around two hundred Linux open source platform vendors globally and they pool together hundreds of applications, placed on top of an open source operating system, marketed through a number of channels and via a number of different business models. It is evident that there is strong competition in this field. Once open source code is *improved* by a firm, by its nature, it is accessible to its customers or even to its competitors. However, due to its complexity of programming, the *inventor* can enjoy advantageous position as the first mover for a span of time. We use a two-period three-stage model where, in each period, firms decide whether to use the open source in the first stage, pursue cost-reducing research in the second stage, and engage in Cournot competition in the third stage.

There are several papers which have dealt with economics of open innovation. For example, Lerner and Tirole (2002) provide a broad discussion of a number of issues, emphasizing career concern. Johnson (2002) and Modica and Aghion (2006) present welfare results and comparative statics using a model of private provision of a public good. Since they use a static model, however, they do not fully capture the main characteristic of the GPL, which is basically "get it for free *now*, pay back *when/if* you succeed." Unlike other papers that use a private provision of a public good type of model, we believe, our model captures the essence of the GPL in a more direct way.

By altering the timing of incentives, open innovation under the GPL has a trade-off between stimulating innovation and promoting disclosure. By using open knowledge, a firm can increase (decrease) its production technology level (unit cost) and therefore its probability of innovation success and of achieving higher profits in that period. Under the GPL, however, any innovative findings that have used open knowledge should also be open knowledge in subsequent periods. This obligation decreases the expected future revenue of the firm. We investigate how such a trade-off influences firms' open source use decision depending on their technology level relative to the open source technology level. If a firm has the same technology level as the open source, it does not have an incentive to use the source. This is true because there is no direct gain from using the source, while there is a potential loss due to the obligation of sharing potential innovation with other firms in the future. On the other hand, if a firm has a lower level of production technology than the open source, it is optimal to use the source. Using the source brings a direct benefit because of the immediate increase in the production technology level. However, using the source also incurs a potential loss in the sense that its innovation success will help its rivals in the second period by the nature of the GPL. We will analyze this trade-off and show that such a potential second period loss is smaller than the loss incurred in the second period from not using the open source in the first period. Hence such a firm will always choose to use the open source. One implication of these, open innovation is developed by lower technology level firms.

We think that it is important to characterize the behavior of a firm when there is open source subject to GPL. This enables us to discuss the welfare aspects of open innovation and compare it to other types of licensing and patent races. This, in turn, will explain why some industries engage in GPL, instead of other types of policies regarding innovation.

Section 2 depicts the model. Section 3 and 4 solves the model backwards. The main results are given in Section 4. In Section 5, we discuss some relevant points and extensions. Section 6 concludes.

2 The Model

There are N firms interacting over two periods. Each firm i produces a good at a firm specific unit cost c_i, which is stochastically determined by investment in cost-reducing innovation. There is also a public production technology, called *open source*, which can produce the good at unit cost c_{os}. In each period, there are three stages: (1) each firm decides whether to adopt the open source or not, (2) each firm invests in cost-reducing innovation, and finally (3) firms compete in quantities in a Cournot fashion. To capture the effect of open source under GPL, we assume that (i) each firm is free to use the open source and (ii) any innovation made by a firm which uses the open source in the first period, will be open source in the second period.

Let k_i denote the production technology level for firm i, in the sense that $c_i = 1 - k_i\delta$ is the unit cost of firm i, where $0 < \delta < \frac{1}{3}$. Let k_{os} denote the production technology level of the open source. Before the first period starts, the public production technology level is $k_{os} = 1$, that is, $c_{os} = 1 - \delta$, and there are M_k firms with unit cost $1 - k\delta$ where $\sum_k M_k = N$ and the initial technology level $k \in \{0, 1\}$.[1] Assume that the initial number of firms for each technology level k, $\{M_k\}_k$, is publicly known.

2.1 First Period

In the first stage, each firm decides whether to use the open source or not. Let $m_i \in \{0, 1\}$ denote the open source use decision for firm i, where 0 stands for *not use*, and 1 for *use*. Call it a *non-user* and *user* firm, respectively. When a firm is indifferent between *using* and *not using* the open source, we assume that it chooses to use it. Denote the production technology level of firm i after its open source use decision with $K(m_i, k_i)$. Then,

$$K(m_i, k_i) = m_i \max(k_i, k_{os}) + (1 - m_i)\, k_i.$$

[1] In our model, no one is doing better than open source in period 1. We will discuss how our model can be extended to more general cases later.

In this chapter, we concentrate on a symmetric equilibrium. That is, we assume that firms under the same conditions (m_i, k_i) make the same decision and write $k_i = k$ when there is no risk of confusion. Let n_k denote the number of the other firms that have cost $1 - k\delta$ on decision after the open source use decision, excluding the firm i's own decision. Using boldface to denote equilibrium number of firms,

$$\begin{aligned} \mathbf{n}_0 &= (1 - \mathbf{m}_0)\, M_0 \\ \mathbf{n}_1 &= M_1 + (1-\alpha)\mathbf{m}_0 M_0 \end{aligned}$$

where m_0 and m_1 are the open source use decisions of the firms with $k_i = 0$ and 1, respectively. Let V_k^m denote the first period expected value of a firm with (m, k) at the end of the first stage. Then, $V_k^1 \geq V_k^0$ if and only if $m = 1$ is preferred to $m = 0$ for each initial $k \in \{0, 1\}$

In the second stage, knowing its own m_i and the equilibrium number of the other firms for each k, $\{\mathbf{n}_k\}_k$, firm i with (m_i, k_i) invests in innovation by picking the probability of success, $p_{k_i}^{m_i}$ at cost $C(p_{k_i}^{m_i}) = \frac{1}{2}(p_{k_i}^{m_i})^2$. That is, the technology is advanced by one level $k_i + 1$ with probability $p_{k_i}^{m_i}$ and it remains in the same level k_i with probability $1 - p_{k_i}^{m_i}$. The problem a firm i with (m_i, k_i) solves will be

$$\max_{p_k^m} \left[p_k^m \{E[\pi_{k+1}] + \overline{W}_k^m\} + (1 - p_k^m)\{E[\pi_k] + \underline{W}_k^m\} - C(p_k^m) \right].$$

$\overline{W}_k^m$ is the expected future payoff when an innovation is successful with (m, k) in period 1. Similarly, $\underline{W}_k^m$ is an expected value for the future when an innovation fails with (m, k) in period 1. Let $E[\pi_k]$ denote the expected profit with technology level k in current period. Let N_k denote the number of firms that have cost $1 - k\delta$ after the innovation realizations.

In the third stage, firms engages in a quantity-setting game a la Cournot. Each firm i decides how much it produces q_{k_i}. At that time, the firm cannot observe other firms' realized technology but has an expected number of firms under a symmetric equilibrium denoted by $\{\mathbf{N}_k\}_k$.[2] Note

[2] $\mathbf{N}_0 = (1 - p_0)\mathbf{n}_0$, $\mathbf{N}_1 = p_0^0 \mathbf{n}_0 + (1 - p_1^0)\mathbf{n}_1^0 + (1 - p_1^1)\mathbf{n}_1^1$, and $\mathbf{N}_2 = p_1^0 \mathbf{n}_1^0 + p_1^1 \mathbf{n}_1^1$. Note that there are two types among firms with $k_i = 1$ after use decision m_i. One is originally having $k_i = 1$ and the other is having $k_i = 1$ thanks to using the open source. To distinguish these two, we use a superscript m_i on n. That is, n_k^m is

that $\sum_k \mathbf{N}_k = N - 1$ since firm i's own is excluded. The expected inverse demand of a firm with k_i is given by $\mathbf{P}_{k_i} \equiv P(\mathbf{Q}_{k_i}) = A - \mathbf{Q}_{k_i}$, where $A > 0$ is sufficiently large, P is the market price, and $\mathbf{Q}_{k_i}$ is the expected total quantity of a firm with k_i. The expected total quantity can be decomposed into two parts, its own quantity, which is known, and expected all other firms' quantities. That is, $\mathbf{Q}_{k_i} = q_{k_i} + \mathbf{Q}_{-i}$ where $\mathbf{Q}_{-i} = \sum_k \mathbf{N}_k q_k$. Under Cournot competition, firm i solves the following problem

$$\max_{q_{k_i}} E[\pi_{k_i}] = (\mathbf{P}_{k_i} - c_{k_i}) q_{k_i}.$$

We assume that firms can observe the realized number of firms, $\{N_k\}_k$ at the end of the stage.

2.2 Second Period

Let X' denote a variable in period 2 when the variable was denoted by X in period 1. For instance, open source use decision in period 2 is denoted by m_i'. In the second period, the first period is essentially repeated but with an important difference. Whenever a firm uses the open source in the first period and succeeds in cost-reducing innovation, the new production technology (the new level of unit cost) becomes free, *in the second period,* to adopt by any other firm. Therefore, whenever there is at least one successful user firm in the first period, the open source is improved by one level, $k_{os}' = \max_{i \in I}\{k_i\}$ where I is a set of user firms. This assumption is motivated by the philosophy of the GPL, which grants the recipients of a computer program the rights of free software definition and ensures this freedom is preserved, even when the work is changed or added to.

3 Equilibrium Analysis: Second Period

To solve this model, we study the subgame perfect equilibrium of the three stage-two period game. Hence we solve the model backwards. But before we do so, we first prove a very useful lemma.

---

the fraction of firms with (k, m). For instance, the fraction of firms with $k = 1$ that has used the open source is given by n_1^1. Then, $\mathbf{n}_1^0 = M_1(1 - \mathbf{m}_1)$ and $\mathbf{n}_1^1 = M_1\mathbf{m}_1 + M_0\mathbf{m}_0$.

Lemma 1 *In period 2, $m'_i = 1$ is a best response for any history of the game, for all i.*

The proof is straightfoward. Any firm with $k'_i = 0$ in the beginning of period 2 will be strictly better off using the open source since open source has a strictly lower its unit cost. Any firm with $k'_i = 1$ will be strictly better off if $k'_{os} = 2$ and will be indifferent if $k'_{os} = 1$. Any firm with $k'_i = 2$ has no benefit from using the open source, so those firms will be indifferent between using and not using the open source. Simply speaking, since there is no future in the last period, there is no negative future effect of using the open source now. Hence all firms at this stage will be weakly better off by using the open source. This lemma allows us to build up backward induction simply. Also note that the lemma is applicable to the final period in any finite game.

3.1 Third Stage: Cournot Competition

Each firm i observes own unit cost k_i and has an expectation of a technology distribution in the market, $\{\mathbf{N}'_k\}_k$.[3] The expected inverse demand with k_i is $\mathbf{P}'_{k_i} \equiv P(\mathbf{Q}'_{k_i}) = A - \mathbf{Q}'_{k_i}$. The expected total quantity can be decomposed; $\mathbf{Q}'_{k_i} = q'_{k_i} + \mathbf{Q}'_{-i}$ where $\mathbf{Q}'_{-i} = \sum_k \mathbf{N}'_k q'_k$ and $\sum_k \mathbf{N}'_k = N - 1$.

By Lemma 1, there is no firm with $k'_i = 0$ after innovation realization in the second period since $k'_{os} = 1$ at least. Thus, $k \in \{1, 2, 3\}$ in the equilibrium when a firm i with k'_i chooses its quantity q'_{k_i}.

$$\max_{q'_{k_i}} E[\pi'_{k_i}] = (\mathbf{P}'_{k_i} - c'_{k_i})q'_{k_i}$$

Then, the equilibrium quantities will be as follows;[4]

$$\begin{aligned}
q'_1 &= \frac{A-1}{N+1} - \frac{\mathbf{N}'_2 + 2\mathbf{N}'_3}{N+1}\frac{\delta}{2} \\
q'_2 &= q'_1 + \frac{\delta}{2} \\
q'_3 &= q'_1 + \delta.
\end{aligned}$$

By using these, we can get the expected profit of a firm with k'_i, $E[\pi'_{k_i}] = (q'_{k_i})^2$ for each k'. It implies that a more advanced firm produces more and gets more profits than less an advanced

[3] Specifically, $\mathbf{R}'_1 = \mathbf{r}'_1(1 - p'_1)$, $\mathbf{R}'_2 = \mathbf{r}'_1(p'_1 + p'_2 - 1) + 1 - p'_2$, and $\mathbf{R}'_3 = (1 - \mathbf{r}'_1)\, p'_2$.

[4] For details, see Appendix 3.A.

firm.

3.2 Second Stage: Investment in Innovation

Similar to the third stage, when firms decide their investment levels, they know their own unit cost but only know a fraction of firms with each k' in equilibrium. Note that at this stage $k'_i \in \{1, 2\}$, that is, $k'_i = 0$ will never be realized. This is because each firm will find it optimal to use the open source in the second period by Lemma 1 and $k'_{os} = 1$ at least. Recall that n'_k means the number of firms that have unit cost $1 - k'\delta$ after the open source use decision but before the investment decision in the second period. Each firm i with unit cost $1 - k\delta$ invests in probabilities of success in innovation, $p'_{k_i} \in (0, 1)$, by maximizing its expected profit. That is,

$$\max_{p'_{k_i}} \left[p'_{k_i} E[\pi'_{k+1}] + (1 - p'_{k_i}) E[\pi'_k] - C(p'_{k_i})\right]$$

By solving the problem, the equilibrium investment of firm i is given by

$$\begin{aligned} p'_k &= E[\pi'_{k+1}] - E[\pi'_k] \\ &= \delta(q'_k + \frac{\delta}{4}). \end{aligned}$$

Recall that $q'_{k_a} > q'_{k_b}$ if and only if $k'_a > k'_b$. Thus, the above result implies that a more advanced firm invests more than a less advanced firm.

3.3 First Stage: Using Open Source

The first stage in period 2 is already discussed in the beginning of this section. And the result is summarized in Lemma 1; any firms prefer to use the open source in the last period.

4 Equilibrium Analysis: First Period

4.1 Third Stage: Cournot Competition

Similar to period 2, a firm i with k_i decides how much it will produce;

$$\max_{q_{k_i}} E[\pi_{k_i}] = (\mathbf{P}_{k_i} - c_{k_i})q_{k_i}$$

A notable difference from period 2 is $k \in \{0, 1, 2\}$ at the third stage. The equilibrium quantity for each k will be

$$\begin{aligned} q_0 &= \frac{A-1}{N+1} - \frac{\mathbf{N}_1 + 2\mathbf{N}_2}{N+1}\frac{\delta}{2} \\ q_1 &= q_0 + \frac{\delta}{2} \\ q_2 &= q_0 + \delta \end{aligned}$$

The same as the second period, the first period equilibrium profit for a firm with unit cost $1 - k\delta$ is $E[\pi_k] = (q_k)^2$ for each $k \in \{0, 1, 2\}$.

4.2 Second Stage: Investment in Innovation

In the investment stage, there are two possible technology levels, $k \in \{0, 1\}$. Each firm invests in probabilities of success in innovation by maximizing its expected profit. However this is more involved than the maximization problem in the investment stage in the second period. Here, on the other hand the investment decision of a firm depends not only on its current unit cost but also on its open source use decision. Hence a firm with unit cost $1 - k\delta$ and open source use decision m chooses p_k^m to maximize the expected profits. That is,

$$\max_{p_k^m} \left[p_k^m \{E[\pi_{k+1}] + \overline{W}_k^m\} + (1 - p_k^m)\{E[\pi_k] + \underline{W}_k^m\} - C(p_k^m) \right]$$

where $\overline{W}_k^m$ and $\overline{W}_k^m$ are the second period expected profit of a firm with (m, k) in period 1 when it has succeeded and failed in period 1, respectively.

Similar to period 2, the optimal investment level in period 1 is given by

$$p_k^m = E[\pi_{k+1}] - E[\pi_k] + [\overline{W}_k^m - \underline{W}_k^m].$$

4.3 First Stage: Using Open Source

In the very first stage, each firm decides whether to use the open source or not. The decision will be based on the initial unit cost. The main trade-off for a firm which has unit cost 1 (that is, $k = 0$) is that using the open source will be beneficial for the first period through lower unit cost, but a potential innovation by such a firm will decrease the unit cost of other firms in the second period, because of the structure imposed by the GPL. Recall that V_k^m is the first period expected value of a firm with (m, k);

$$V_k^m = p_k^m[E[\pi_{k+1}] + \overline{W}_k^m] + (1 - p_k^m)\,[E[\pi_k] + \underline{W}_k^m] - C(p_k^m)$$

The optimal decision for a firm with initial unit cost $1 - \delta$, that is, $k = 1$, is not to use the open source. Any firm with unit cost 1, that is, $k = 0$, will choose to use the open source. This is summarized, respectively, in the two propositions below. But first we provide some useful identities which we will use in the proof of the propositions, after introducing several notations.

Let W_k^{inf} denote the second period expected profit for a firm with k when the firm is one of *the least advanced*, that is, when there exists at least one other firm with $\widehat{k} > k$ and no other firms with $\widetilde{k} < k$ before the second period investment decisions are made. W_k^{sup} is for the case when the firm is one of *the most advanced*, that is, there exists at least one other firm with $\widetilde{k} < k$ and no other firm with $\widehat{k} > k$. Finally, W_k^{equ} implies the expected profit when every firm has k.

Now consider a case when a firm fails to innovate in the first period. In such a case, the firm's *use* decision does not affect its expected second period profits. To see this, first note that the firm's *use* decision does not improve the open source in the second period. Secondly, the firm will use the open source in the second period by Lemma 1, so its expected future profit only depends on other firms' behavior. Therefore, the second period expected profit of such a firm is independent of its *use* decision. Note that this is true for any k. That is, for any (m, k),

$$\underline{W} = \underline{W}_k^m = \eta[\tau W_1^{equ} + (1-\tau)W_1^{inf}] + (1-\eta)W_2^{equ} \tag{1}$$

where η is the probability that every *other* user firm fails to innovate in the first period, and τ_α is the probability that all *other* non-user firms initially with $k = 1$ fail to innovate in the first period.

With probability $1-\eta$, at least one user firm succeeds. Then, the open source will be $k'_{os} = 2$, so every firm will have $k' = 2$. With probability η, no user firm makes a success so $k'_{os} = 1$. In such a case, either every firm has $k' = 1$, or there is at least one other firm with $k' = 2$ by Lemma 1. The former has probability τ, and the latter has probability $1 - \tau$. Hence the expression above.

Note that a non-user firm with initial $k = 0$ can not improve the source even when it succeeds in the first period. Hence we have the identity below.

$$\overline{W}_0^0 = \underline{W} \tag{2}$$

A non-user firm with initial $k = 1$ will have $k' = 2$ if it succeeds. In this case, when none of the user firms succeed, the open source stays at $k'_{os} = 1$, so its profit is W_2^{sup}. Otherwise, every firm will have $k = 2$ by Lemma 1 so its profit will be W_2^{equ}. Hence the firm will earn W_2^{sup} with probability η, and W_2^{equ} with probability $1 - \eta$. Therefore we have the following identity.

$$\overline{W}_1^0 = \eta W_2^{sup} + (1-\eta)W_2^{equ} \tag{3}$$

Finally, if a user firm succeeds, then the source will be $k'_{os} = 2$, so does every other firm. Therefore the firm will get W_2^{equ}, regardless of its initial cost. Hence the expression

$$\overline{W}_1^1 = \overline{W}_0^1 = W_2^{equ} \tag{4}$$

Now we can prove our first result.

Proposition 1 *A firm with the same technology level as the open source does not use the open source in period 1.*

Proof. See the Appendix B. ■

The intuition for this result is that whenever a firm starts the game with the same unit cost as the open source has, there is no direct benefit from using the open source. However, using the open source makes the firm obliged to share its potential first period innovation, with other firms, in the second period, who choose to use the open source in the second period. This removes any potential cost advantage the firm could have in the second period quantity setting game. Hence any such firm will avoid using the open source.

Our second result says that any firm that produces the good at a higher unit cost than the open source will choose to use the open source even though GPL makes the firm obliged to share its potential innovation in the first period with the firms in the second period.

Proposition 2 *A firm with lower technology level than the open source chooses to use the open source in period 1.*

Proof. See the Appendix B. ■

The intuition for this result is as follows. For a firm with a lower technology level (that is, a higher unit cost) than the open source, it is clear that using the open source directly improves the firm's production technology hence its expected profit in period 1. Now compare the second period expected profit from between using and not using the open source. First consider a case where the first period innovation fails. Then, by (1), the second period expected profit is the same between a user firm and a non-user firm. Now consider when the first period innovation succeeds. The second period expected profit of a user firm is V_2^{equ} since every other firm will share its innovation. For a non-user firm, its technology level in the beginning of period 2 is $k' = 1$. Hence, the second period expected profit of a non-user firm is at best V_2^{equ}, or possibly V_1^{inf} when all user firms has failed and some non-user firm with initial $k = 1$ make a success. In sum, for a firm with a lower initial technology than the open source, using the open source is beneficial for the second period in expectation as well as the first period. Therefore, the incentive to use the open source dominates the incentive not to use it.

5 Discussion

In this section we discuss several relevant points and extensions. One of the most important questions is whether the open innovation under GPL improves social welfare of the economy and, if it does, by how much. Based on firm's optimal use decision, our model provides a way to measure the welfare gain and loss from open source. The welfare gain comes from stronger competition in the production stage than without open source case. In the innovation race, open source helps firms that are *behind* to keep chasing firms that are *ahead*. On the other hand, it increases the likelihood that advanced firms are caught relative to the situation without open source. Then, there will be larger consumer surplus under the open source than without it. However, firms have investment costs as well as production costs. A stronger competition under the open source may lead both advanced firms and lagged firms invest more than the socially optimal level. We can measure how much social welfare gain comes from the open source as $SW = SW(k_{os} = 1) - SW(k_{os} = 0)$ where $SW(k_{os}) = CS(k_{os}) + PS(k_{os})$, CS is consumer surplus, and PS is producer surplus;

$$
\begin{aligned}
CS(k_{os}) &= \frac{1}{2}(A-P)Q + \frac{1}{2}(A-P')Q' \\
&= \frac{1}{2}(A-(A-Q))Q + \frac{1}{2}(A-(A-Q^{'}))Q' \\
&= \frac{1}{2}(Q^2 + Q'^2) \\
&= \frac{1}{2}[(\textstyle\sum_{k=0}^{2} N_k q_k)^2 + (\textstyle\sum_{k=1}^{3} N'_k q'_k)^2]
\end{aligned}
$$

and

$$
\begin{aligned}
PS(k_{os}) &= \textstyle\sum_{k=0}^{2} N_k \left[\pi_k - C(p_k)\right] + \sum_{k=1}^{3} N'_k \left[\pi'_k - C(p'_k)\right] \\
&= \textstyle\sum_{k=0}^{2} N_k \left[(q_k)^2 - \delta(q_k + \dfrac{\delta}{4})\right] + \sum_{k=1}^{3} N'_k \left[(q'_k)^2 - \delta(q'_k + \dfrac{\delta}{4})\right].
\end{aligned}
$$

Another interesting question related with welfare analysis in open innovation is whether the social welfare from GPL dominates that from traditional licenses or patent systems. Our model allows one to quantitatively do welfare comparisons and to determine the conditions for a welfare gain.

The initial unit cost distribution we have assumed is a specific one; k_i is equal to either 1 or 0. Instead, we can assume a more general distribution. When the initial $k \in \{0, 1, \cdots, 2M\}$ and $k_{os} = M$, we can describe firms with lower and higher technology levels than the open source accordingly. Under this assumption, the idea behind our results will still be valid, that is, the firms that have a higher unit cost than the source will use the source, the firms that have the same or lower unit cost will choose not to use the source. Then, we can extend our model to more than two periods easily. Analyzing a longer horizon, we can understand the evolution of the open source technology together with use decisions of the firms over time.

6 Conclusion

In this paper, we analyzed a simple model of innovation in cost reduction with an open source production technology present for the firms to freely use. We assumed, in the spirit of the GPL, that whenever a user succeeds in cost reduction innovation, it has to share this new technology with other users. Because of this aspect of the GPL, we used a dynamic model with two periods. We characterized the optimal open source use decision of a firm as a function of its technology level relative to the open source. A firm that has the same technology as the open source finds it optimal not to use the open source. A firm that has a lower production technology level finds it optimal to use the source.

REFERENCES

1. Aghion, Philippe, Salvatore Modica, 2006. "Open Source without Free-Riding," mimeo.

2. Aoki, Reiko, 1991. R&D Competition for Product Innovation: An Endless Race, The American Economic Review, 81:252-256.

3. Johnson, Justin, 2002, "Open Source Software: Private Provision of a Public Good." Journal of Economics and Management Strategy, 11:637-662.

4. Lerner, Josh and Jean Tirole, 2002. "The Simple Economics of Open Source," Journal of Industrial Economics, 52:197-234.

5. Merton, Robert M., 1968. The Matthew Effect in Science, Science, 159:56-63

7 Appendix A

7.1 Cournot Competition

Recall that the expected inverse market demand with knowing its own technology level k_i is given by $\mathbf{P}_{k_i} \equiv P(\mathbf{Q}_{k_i}) = A - \mathbf{Q}_{k_i}$. And $\mathbf{Q}_{k_i} = q_{k_i} + \mathbf{Q}_{-i}$ where $\mathbf{Q}_{-i} = (\mathbf{N}_0 q_0 + \mathbf{N}_1 q_1 + \mathbf{N}_2 q_2)$ and $\mathbf{N}_0 + \mathbf{N}_1 + \mathbf{N}_2 = N - 1$. Under Cournot competition, firm i solves the following problem

$$\max_{q_{k_i}} E[\pi_{k_i}] = (\mathbf{P}_{k_i} - c_{k_i}) q_{k_i}$$

From the first order condition,

$$\begin{aligned}
\frac{\partial E[\pi_{k_i}]}{\partial q_{k_i}} &= P(\mathbf{Q}_{k_i}) - c_{k_i} + \frac{\partial P(\mathbf{Q}_{k_i})}{\partial q_{k_i}} q_{k_i} \\
&= A - \mathbf{Q}_{k_i} - c_{k_i} - q_{k_i} \\
&= A - c_{k_i} - \mathbf{Q}_{-i} - 2q_{k_i}
\end{aligned}$$

since $\frac{\partial \mathbf{Q}_{k_i}}{\partial q_{k_i}} = \frac{\partial (q_{k_i}+\mathbf{Q}_{-i})}{\partial q_{k_i}} = 1$. From the fact $c_{k_i} = 1 - k_i\delta$, we have the following equilibrium condition for each k_i;

$$2q_{k_i} = A - (1 - k_i\delta) - (\mathbf{N}_0 q_0 + \mathbf{N}_1 q_1 + \mathbf{N}_2 q_2)$$

It is easy to see $q_{k_i+1} = q_0 + \frac{\delta}{2}k_i$. Plugging each k_i into the above equation, we get the equilibrium quantity at $k_i = 0$,

$$q_0 = \frac{A-1}{N+1} - \frac{\mathbf{N}_1 + 2\mathbf{N}_2}{N+1}\frac{\delta}{2}.$$

Now we calculate the equilibrium profit levels. First, note that

$$\begin{aligned}
\mathbf{Q}_{k_i} &= q_{k_i} + N(\mathbf{N}_0 q_0 + \mathbf{N}_1 q_1 + \mathbf{N}_2 q_2) \\
&= q_0 + \frac{\delta}{2}k_i + (\mathbf{N}_0 + \mathbf{N}_1 + \mathbf{N}_2)q_0 + (\mathbf{N}_1 + 2\mathbf{N}_2)\frac{\delta}{2} \\
&= q_0 + \frac{\delta}{2}k_i + (N-1)q_0 + (\mathbf{N}_1 + 2\mathbf{N}_2)\frac{\delta}{2} \\
&= \frac{\delta}{2}k_i + Nq_0 + (\mathbf{N}_1 + 2\mathbf{N}_2)\frac{\delta}{2} \\
&= \frac{\delta}{2}k_i + N\left(\frac{A-1}{N+1} - \frac{N}{N+1}(\mathbf{N}_1 + 2\mathbf{N}_2)\frac{\delta}{2} + (\mathbf{N}_1 + 2\mathbf{N}_2)\frac{\delta}{2}\right) \\
&= \frac{\delta}{2}k_i + \frac{N}{N+1}\left(A - 1 + (\mathbf{N}_1 + 2\mathbf{N}_2)\frac{\delta}{2}\right)
\end{aligned}$$

And

$$\begin{aligned}
\mathbf{P}_{k_i} &= A - \mathbf{Q}_{k_i} = \frac{A-1}{N+1} - \frac{\mathbf{N}_1 + 2\mathbf{N}_2}{N+1}\frac{\delta}{2} + 1 - \frac{\delta}{2}k_i \\
&= q_0 + 1 - \frac{\delta}{2}k_i
\end{aligned}$$

Hence we get

$$\boldsymbol{\pi}_{k_i} = (\mathbf{P}_{k_i} - c_{k_i})q_{k_i} = \left[q_0 + 1 - \frac{\delta}{2}k_i - (1 - \frac{\delta}{2}k_i)\right]q_{k_i} = (q_{k_i})^2$$

The second period Cournot competition yields similar expressions.

$$
\begin{aligned}
q_1' &= \frac{A-1}{N+1} - \frac{\mathbf{N}_2' + 2\mathbf{N}_3'}{N+1}\frac{\delta}{2} \\
q_2' &= q_1' + \frac{\delta}{2} \\
q_3' &= q_1' + \delta
\end{aligned}
$$

and $\pi_k' = (q_k')^2$, where $k \in \{1, 2, 3\}$.

8 Appendix B

Proof of Proposition 1. Given $c_i = 1 - \delta$, the optimality of the investment levels p_1^0 and p_1^1 implies

$$
\frac{\partial V_1^0}{\partial p_1^0} = [E[\pi_2] - E[\pi_1]] + [\overline{W}_1^0 - \underline{W}_1^0] - C'(p_1^0) = 0, \text{ that is, } p_1^0 = E[\pi_2] - E[\pi_1]] + [\overline{W}_1^0 - \underline{W}_1^0]
$$

$$
\frac{\partial V_1^1}{\partial p_1^1} = [E[\pi_2] - E[\pi_1]] + [\overline{W}_1^1 - \underline{W}_1^1] - C'(p_1^1) = 0, \text{ that is, } p_1^1 = [E[\pi_2] - E[\pi_1]] + [\overline{W}_1^1 - \underline{W}_1^1]
$$

By (1), $\underline{W}_1^0 = \underline{W}_1^1$, we have $p_1^0 - p_1^1 = \overline{W}_1^0 - \overline{W}_1^1$ and $p_1^0 + p_1^1 = 2(E[\pi_2] - E[\pi_1] - \underline{W}_1^0) + \overline{W}_1^0 + \overline{W}_1^1$. Now we can calculate $V_1^0 - V_1^1$.

$$
\begin{aligned}
V_1^0 - V_1^1 &= \left(p_1^0 - p_1^1\right)\left[E[\pi_2] - E[\pi_1] - \underline{W}_1^0\right] + p_1^0\overline{W}_1^0 - p_1^1\overline{W}_1^1 - \frac{1}{2}\left(p_1^0 - p_1^1\right)\left(p_1^0 + p_1^1\right) \\
&= \left(p_1^0 - p_1^1\right)\left[E[\pi_2] - E[\pi_1] - \underline{W}_1^0 - \frac{1}{2}\left(p_1^0 + p_1^1\right)\right] + p_1^0\overline{W}_1^0 - p_1^1\overline{W}_1^1
\end{aligned}
$$

$$
\begin{aligned}
&= \left(p_1^0 - p_1^1\right)\left[E[\pi_2] - E[\pi_1] - \underline{W}_1^0 - (E[\pi_2] - E[\pi_1] - \underline{W}_1^0) - \frac{1}{2}(\overline{W}_1^0 + \overline{W}_1^1)\right] + p_1^0\overline{W}_1^0 - p_1^1\overline{W}_1^1 \\
&= \left(p_1^0 - p_1^1\right)\left[-\frac{1}{2}(\overline{W}_1^0 + \overline{W}_1^1)\right] + p_1^0\overline{W}_1^0 - p_1^1\overline{W}_1^1 \\
&= \frac{1}{2}(\overline{W}_1^0 - \overline{W}_1^1)((p_1^0 + p_1^1))
\end{aligned}
$$

Hence, $V_1^0 > V_1^1$ if and only if $\overline{W}_1^0 > \overline{W}_1^1$. That is, $V_1^0 > V_1^1$ if and only if $\eta W_2^{sup} + (1-\eta) W_2^{equ} > W_2^{equ}$, by (3) and (4). We can write,

$$\begin{aligned} W_2^{sup} &= \max_{p_2'} \left[p_2' E[\pi_3' | \mathbf{n}_2' < 1] + \left(1 - p_2'\right) E[\pi_2' | \mathbf{n}_2' < 1] - C(p_2') \right] \\ W_2^{equ} &= \max_{p_2'} \left[p_2' E[\pi_3' | \mathbf{n}_2' = 1] + \left(1 - p_2'\right) E[\pi_2' | \mathbf{n}_2' = 1] - C(p_2') \right] \end{aligned}$$

Recall that $\pi_k' = (q_k')^2 = \left[\frac{A-1}{N+1} - \frac{\mathbf{N}_2' + 2\mathbf{N}_3'}{N+1} \frac{\delta}{2} + (k-1) \frac{\delta}{2} \right]^2$. Since $\mathbf{N}_2' + 2\mathbf{N}_3' = p_2' \mathbf{n}_2' + (1 - r_2') p_1' + \mathbf{n}_2'$, we have $\frac{\partial \left(\mathbf{N}_2' + 2\mathbf{N}_3' \right)}{\partial \mathbf{n}_2'} = p_2' - p_1' + 1 > 0$ because $p_2' > p_1'$. Hence, π_k' is strictly decreasing in $\mathbf{n}_2'$ since $\frac{\partial \pi_k'}{\partial \left(\mathbf{N}_2' + 2\mathbf{N}_3' \right)} < 0$. Therefore $E[\pi_k' | \mathbf{n}_2' < 1] > E[\pi_k' | \mathbf{n}_2' = 1]$ for $k \in \{2, 3\}$. This implies that $p_2' E[\pi_3' | \mathbf{n}_2' < 1] + (1 - p_2') E[\pi_2' | \mathbf{n}_2' < 1] > p_2' E[\pi_3' | \mathbf{n}_2' = 1] + (1 - p_2') E[\pi_2' | \mathbf{n}_2' = 1]$ for any p_2'. Since both W_2^{sup} and W_2^{equ} are strictly increasing in p_2', we have $W_2^{sup} > W_2^{equ}$. Therefore $V_1^0 > V_1^1$. ■

Proof of Proposition 2. We already have

$$\begin{aligned} V_0^0 &= p_0^0 [E[\pi_1] + \overline{W}_0^0] + \left(1 - p_0^0\right) [E[\pi_0] + \underline{W}_0^0] - C(p_0^0) \\ V_0^1 &= p_0^1 [E[\pi_2] + \overline{W}_0^1] + \left(1 - p_0^1\right) [E[\pi_1] + \underline{W}_0^1] - C(p_0^1) \end{aligned}$$

Rearranging terms we can write,

$$\begin{aligned} V_0^0 &= p_0^0 (E[\pi_1] - E[\pi_0]) + E[\pi_0] + (p_0^0 \overline{W}_0^0 + (1 - p_0^0) \underline{W}_0^0) - \frac{1}{2} (p_0^0)^2 \\ V_0^1 &= p_0^1 (E[\pi_2] - E[\pi_1]) + E[\pi_1] + (p_0^1 \overline{W}_0^1 + (1 - p_0^1) \underline{W}_0^1) - \frac{1}{2} (p_0^1)^2 \end{aligned}$$

First order conditions are

$$\begin{aligned} p_0^0 &= E[\pi_1] - E[\pi_0] + [\overline{W}_0^0 - \underline{W}_0^0] \\ p_0^1 &= E[\pi_2] - E[\pi_1] + [\overline{W}_0^1 - \underline{W}_0^1] \end{aligned}$$

Plugging these into the expected profit expressions, we get,

$$\begin{aligned}
V_0^0 &= (E[\pi_1] - E[\pi_0] + \overline{W}_0^0 - \underline{W}_0^0)(E[\pi_1] - E[\pi_0]) + E[\pi_0] \\
&\quad + (E[\pi_1] - E[\pi_0] + \overline{W}_0^0 - \underline{W}_0^0)(\overline{W}_0^0 - \underline{W}_0^0) + \underline{W}_0^0 - \frac{1}{2}(p_0^0)^2 \\
&= (E[\pi_1] - E[\pi_0])^2 + 2(\overline{W}_0^0 - \underline{W}_0^0)(E[\pi_1] - E[\pi_0]) + (\overline{W}_0^0 - \underline{W}_0^0)^2 \\
&\quad + E[\pi_0] + \underline{W}_0^0 - \frac{1}{2}(p_0^0)^2 \\
&= (E[\pi_1] - E[\pi_0] + \overline{W}_0^0 - \underline{W}_0^0)^2 + E[\pi_0] + \underline{W}_0^0 - \frac{1}{2}(p_0^0)^2 \\
&= \frac{1}{2}(p_0^0)^2 + E[\pi_0] + \underline{W}_0^0
\end{aligned}$$

Similarly we have,

$$V_0^1 = \frac{1}{2}(p_0^1)^2 + E[\pi_1] + \underline{W}_0^1$$

Hence,

$$\begin{aligned}
V_0^1 - V_0^0 &= \frac{1}{2}(p_0^1)^2 + E[\pi_1] + \underline{W}_0^1 - (\frac{1}{2}(p_0^0)^2 + E[\pi_0] + \underline{W}_0^0) \\
&= \frac{1}{2}[(p_0^1)^2 - (p_0^0)^2] + E[\pi_1] - E[\pi_0] + \underline{W}_0^1 - \underline{W}_0^0 \\
&= \frac{1}{2}[(p_0^1)^2 - (p_0^0)^2] + E[\pi_1] - E[\pi_0] \text{ by (1).}
\end{aligned}$$

Since $E[\pi_1] > E[\pi_0]$, it suffices to show $p_0^1 > p_0^0$. Recall that $\pi_k = (q_k)^2 = (q_0 + k\frac{\delta}{2})^2$ where $q_0 = \frac{A-1}{N+1} - \frac{\mathbf{N}_1 + 2\mathbf{N}_2}{N+1}\frac{\delta}{2}$. Then $E[\pi_k] = E[q_k]^2$ where

$$E[q_k] = \frac{A-1}{N+1} - \frac{\mathbf{N}_1 + 2\mathbf{N}_2}{N+1}\frac{\delta}{2} + k\frac{\delta}{2} = E[q_0] + k\frac{\delta}{2} \tag{5}$$

Then,

$$
\begin{aligned}
p_0^0 &= E[\pi_1] - E[\pi_0] + [\overline{W}_0^0 - \underline{W}_0^0] \\
&= E[q_1]^2 - E[q_0]^2 \qquad \text{by (2) and (5).} \\
&= (E[q_0] + \frac{\delta}{2})^2 - E[q_0]^2 \\
&= \delta(E[q_0] + \frac{\delta}{4}) \\
p_0^1 &= E[\pi_2] - E[\pi_1] + [\overline{W}_0^1 - \underline{W}_0^1] \\
&= E[q_2]^2 - E[q_1]^2 + W_2^{equ} - \underline{W}_0^1 \qquad \text{by (4).} \\
&= (E[q_0] + \delta)^2 - (E[q_0] + \frac{\delta}{2})^2 + W_2^{equ} - \underline{W}_0^1 \\
&= \delta(E[q_0] + 3\frac{\delta}{4}) + W_2^{equ} - \underline{W}_0^1
\end{aligned}
$$

Note that, $W_2^{equ} > \eta[\tau W_1^{equ} + (1-\tau)W_1^{inf}] + (1-\eta)W_2^{equ}$, since $W_2^{equ} > W_1^{equ} = \max\{W_1^{equ}, W_1^{inf}\}$. To see this,

$$
\begin{aligned}
W_1^{inf} &= \max_{p_2'} \left[p_2' E[\pi_2' | \mathbf{r}_1' < 1] + \left(1 - p_2'\right) E[\pi_1' | \mathbf{n}_1' < 1] - C(p_2') \right] \\
V_1 &= \max_{p_2'} \left[p_2' E[\pi_2' | \mathbf{r}_1' = 1] + \left(1 - p_2'\right) E[\pi_1' | \mathbf{n}_1' = 1] - C(p_2') \right] \\
W_2^{equ} &= \max_{p_2'} \left[p_2' E[\pi_3' | \mathbf{r}_2' = 1] + \left(1 - p_2'\right) E[\pi_2' | \mathbf{n}_2' = 1] - C(p_2') \right]
\end{aligned}
$$

Recall that $\pi_k' = (q_k')^2 = \left[\frac{A-1}{N+1} - \frac{\mathbf{N}_2' + 2\mathbf{N}_3'}{N+1} \frac{\delta}{2} + (k-1)\frac{\delta}{2} \right]^2$. Since $\mathbf{N}_2' + 2\mathbf{N}_3' = p_2'(1 - \mathbf{n}_1') + \mathbf{n}_1' p_1' + 1 - \mathbf{n}_1'$, we have $\frac{\partial\left(\mathbf{N}_2' + 2\mathbf{N}_3'\right)}{\partial \mathbf{n}_1'} = p_1' - p_2' - 1 < 0$ because $p_2' > p_1'$. Hence, π_k' is strictly increasing in $\mathbf{n}_1'$ since $\frac{\partial \pi_k'}{\partial\left(\mathbf{R}_2' + 2\mathbf{R}_3'\right)} < 0$. Therefore $E[\pi_k' | \mathbf{n}_1' < 1] < E[\pi_k' | \mathbf{n}_1' = 1]$ for $k \in \{1, 2\}$. This implies that $p_2' E[\pi_2' | \mathbf{n}_1' < 1] + (1 - p_2') E[\pi_1' | \mathbf{n}_1' < 1] < [p_2' E[\pi_2' | \mathbf{n}_1' = 1] + (1 - p_2') E[\pi_1' | \mathbf{n}_1' = 1] - C(p_2')]$ for any p_2'. Since both W_1^{inf} and W_1^{equ} are strictly increasing in p_2', we have $W_1^{equ} > W_1^{inf}$. Since $E[\pi_{k+1}' | \mathbf{n}_1' = 1] > E[\pi_k' | \mathbf{n}_1' = 1]$,it directly implies that $W_2^{equ} > W_1^{equ}$. Since $W_1^{equ} > W_1^{inf}$,we conclude $W_2^{equ} > W_1^{equ} = \max\{W_1^{equ}, W_1^{inf}\}$. Hence $W_2^{equ} - \underline{W}_0^1 > 0$, which implies $p_0^1 > p_0^0$. Therefore $V_0^1 > V_0^0$. ■