The Rationale Behind P2P Network Users

Brian M. Choi

Faculty Advisor: Professor Huseyin Yildirim

Honors Thesis submitted in partial fulfillment of the requirements for Graduation with Distinction in Economics in Trinity College of Duke University

Duke University
Durham, North Carolina
2008

Acknowledgements

The author would like to thank Dr. Yildirim for his invaluable guidance and advice. He would also like to thank Dr. Connolly and econ 198S, 199S classmates for their feedback and constructive criticisms.

Abstract

An online distribution of copyrighted materials through P2P network is one of major legal issues today. While several P2P users have been fined severely for a violation of a copyright law, an illegal file sharing within the P2P network still continues to thrive, and the effectiveness of the punishment against P2P users is often questioned. In this paper, I examine why P2P networks prosper in spite of government’s heavy punishment schemes. Then I model the optimal decision of the P2P user as a non-cooperative game and find out how effectively legal sanction reduces an illegal online distribution of copyrighted materials.

Section I

Introduction: Since the emergence of a P2P network, the online distribution of copyrighted contents has surged. A lot of restrictive measures are taken toward P2P network to protect the copyrighted materials from illegal distribution, but the demand to access P2P network is still high, and the circulation of the copyrighted contents is still not kept in check. As evidence, in 2004, there were more than 8.5 million users utilizing emule or KaZaA to offer and download movies illicitly at any given time (Graham 2004).

The way file sharing occurs through P2P network is quite simple. Once P2P users upload a certain file, as long as those uploaders stay connected to the network (the activity of uploading alone does not guarantee a successful download. Uploaders must also stay "connected" to the network in order for others to download the files that they uploaded), the file becomes available for most, if not all, P2P users, to download. Compared to the number of downloaders in the network, the number of uploaders is really small. Nearly half of all search responses come from the only top 1% of sharing nodes (Ma, Lee, Lui, Yau, 2001). However, in huge P2P networks such as emule, the size of the network compensates for a small portion of uploaders. For instance, if the network has over 100 million users, 1% of the P2P users would still be well over a million people, and they can upload a plethora of files.

Downloading or uploading a copyrighted file is a clear violation of copyright and can make one face a significant amount of fine. What might be P2P users' motive for participating in an illegal file sharing? There have been scholarly studies of incentives for P2P users. The motive behind downloaders' action seems quite intuitive. They want to own an item they want for free and fast. However, the rationale behind the uploaders is not intuitively clear since uploading activity is not pre-requisite for downloading in many P2P networks. Although many P2P networks developed mechanisms that reward generous uploaders, the rewards are generally deemed insufficient to dominate the risk of "punishment" for the violation of copyrights. The motives for P2P users will be discussed in literature review section, and the explanation heavily relies on the psychological aspects; it is also consistent with the basic economic assumption that all the agents are rational.

The main aims of the paper include 1)modeling the optimal decision of the P2P user as a non-cooperative game in which each user's decision depends on other peers' decisions, though in absence of an explicit coordination, and 2)calculating efficiency of one of the most popularly-employed punishment mechanisms by the government, systematically punishing users who illicitly supply copyrighted contents.

Overall, the paper will be divided into four sections: next section will review literature dealt with motives for P2P users, and using the motives discussed in section II, I will set up a relevant payoff function for P2P users in Section III. Furthermore, the theoretical framework for a P2P network of multiple users, the extent to which the legal sanction affects P2P network will also be explained in section III, and section IV will be the conclusion.

Section II

Literature Review: The previous literature on this issue usually focuses on finding an efficient way of protecting copyrighted materials from illegal distribution or how illegal distribution affects the actual sales of the materials. Only a small portion focuses on the rationale behind the actions of downloaders and uploaders.

Most of the literature has similar explanations for the downloaders. Utility they garner from obtaining copyrighted items free and fast dominates the expected cost of downloading, which is a potential legal sanction. Also, an item uploaded on the P2P network can be downloaded an infinite number of times. This “absence of consumption rivalry and the missing excludability from consumption” (Ostrom, Gardner, & Walker, 1994) also increase individuals’ tendency to download. However, the explanation for the uploaders is much more varied.

The paper by Jan U. Becker and Michel Clement, titled “Dynamics of Illegal Participation in Peer-to-Peer Networks – Why Do People Illegally Share Media Files,” published in *Journal of Media Economics* in 2006, extensively surveys the rationale behind uploading a copyrighted material, addressing several factors that contribute to such behavior, and then explain how the P2P network itself is sustained although most of the P2P users only download the files rather than upload them.

Since the P2P system is, by definition, a peer-to-peer network, the optimal decision is not only determined by the user’s own actions but also by the actions of other peers. To calculate the optimal decision for the user, it is imperative to set up a utility function for the P2P user and compare the utilities individuals derive from different decisions. According to the article, individuals derive utility from collecting media files even without consuming them, and thus, the utility function is positively correlated with the number of media files that the individual owns, not the number of the files that he or she consumes. Also, as rational agents, they are aware of the fact that they have to take a risk of a legal sanction when they upload the materials, so their payoff function has a negative correlation with the amount of risk they feel they have to take when upload the files. The size of the network overall did not affect the legal fear as much as the frequency with which individuals access the network. Thus, the optimal strategy, since utility is directly related to downloaded files and inversely related to the risk of getting caught, would be dictated by an incentive to free-ride on the uploaders.

If nobody uploads a file; in other words, if everybody chooses to employ an optimal strategy, then surely, the P2P system will not sustain. However, in reality, P2P networks are booming. There are a plenty of files available in the P2P network, meaning there are always uploaders who satisfy the demand of the free-riders. If downloading as many files as possible while not uploading any files is the optimal strategy, it is not sensible why there are always people who upload files to the P2P network, why they choose not to employ the optimal strategy; it goes against the fundamental economic assumption that all the agents are rational.

However, the paper demonstrates that the payoff for the uploaders is indeed great enough to overcome the fear of a legal sanction when psychological aspects are taken into consideration.

One possible explanation that the paper suggests is that the P2P system serves as a medium through which the uploaders socially interact with the other people online. They feel that their contribution, in this case, uploading copyrighted materials, is appreciated by other P2P users, and this motivates them to continue to upload. Another explanation may be that the uploaders feel their uploading files will motivate others to upload as well, so simply, whenever they upload they expect a favor in return. Therefore, their payoff function includes other factors such as altruism and amount of attention they feel they receive when uploading a file. It is not a function of just two variables: files and risk.

Utilizing data obtained from a sample of 370 people who responded to inquiries of their motives for accessing the P2P network, Becker and Clement empirically found a positive correlation with “perceived importance,” “altruism,” “reciprocity,” and a negative correlation with “legal fear,” confirming their hypotheses.

Becker and Clement also did an additional survey in an attempt to find out if there exist other factors that may have a noticeable impact on the user’s decision to upload files. The survey included questions about how long and how often each respondent has been accessing the P2P network, how much of the downloaded files they share with other P2P users, and motives for uploading files. The results obtained from the survey demonstrate that although most of the P2P users are medium sharers, they do not access the P2P network as often as the generous sharers. As the users access the P2P network for a longer period of time, they develop a loyalty to the P2P network, and feel that they have to support it by uploading more files. Although they fear the consequences of their action, their loyalty overcomes their fear of punishment.

Including all those motives for uploading would severely complicate the model. Instead, I will assume that the only motivation for uploaders in my model is to have access to the network and others’ files, and since medium sharers constitute the majority of P2P users and therefore have most impact on the P2P network. I will limit my model to medium sharers since according to Becker and Clement’s paper, generous sharers will always continue to upload files regardless of the magnitude of the legal fear.

Section III

Theoretical Framework I: One of the basic assumptions in a game-theoretical model is that all economic agents are rational. As a rational economic agent, each P2P user acts in a way that maximizes his or her utility and decides whether to buy or to download the item that he or she wants subject to a certain payoff function. In my model, for any given item, every individual has three actions to choose from, 1) purchasing the item at the (legal) market price, 2) downloading the item for free , and 3) staying idle (not doing anything). As a utility-maximizing economic agent, the individual will choose to do what will generate the highest level of utility among those three options.

The variables in my model include:

α = probability that an individual successfully downloads the file from the P2P network within the expected amount of time (For the sake of simplicity, the expected amount of waiting time is constant and same for everybody and denoted by t)

ν = how much the individual values the file

δ = discount factor between 0 and 1

γ = exogenous probability that an individual uploads the file that others want

ε = legal fear

e = probability that an individual will enter the P2P network

p = market price of the item

N = population size

Let U_t be the average payoff for a person who has had the item for the past t-1 periods. Here, a period is some time unit, which may be measured in minutes. Given that the person's choices are exactly the same in each period following an unsuccessful search online, we drop the time index in the analysis below. As stated above, an individual will choose one of the three actions in period t to maximize U; so

$$U = \max\{u(\text{download}), u(\text{purchase}), u(\text{stay idle})\} \tag{*}$$

in which u(x) denotes utility one garners from choosing a strategy x. Therefore, in order to define a payoff function for an individual, it is necessary to find u(download), u(purchase), u(stay idle) in terms of defined parameters.

If one decides to purchase an item, the price of the item will be deducted from how much he or she actually values the item when calculating how much utility one derives from it, so u(download) will be a linearly decreasing function of p, and a linearly increasing function of ν. Thus,

$$u(\text{purchase}) = \nu - p. \tag{1}$$

Finding u(download) is more subtle. Whether the individual will successfully download the item or not will be subject to the probability, α, and if the probability is 1, then he will receive ν without having to pay price. If the probability is smaller than 1, then with α probability (to be determined in equilibrium), he will download the complete item and receive ν, and with $(1 - \alpha)$ probability, he will not be able to download the item successfully and face the three actions again. Then he will have to wait for more time than he expected, and in that case, he will have to enjoy the item in later time that he expected to enjoy at that given moment, so his utility will be reduced by a factor of δ, a discount factor smaller than 1. Also, regardless of the value of α, since using P2P network to download the copyrighted material is illegal and can lead to a legal sanction, a legal fear, ε will always be deducted from his utility. Therefore,

$$u(\text{download}) = \alpha\nu + (1 - \alpha)\delta U - \varepsilon \tag{2}$$

We normalize and set

u(stay idle) = 0. (3)

If an individual chooses not to do anything, he will receive neither any utility nor any disutility.

Combining (1), (2), (3) with (*), I get the following recursive formula for the person's search problem:

U = max{ ν – p, α ν + (1 – α)δU – ε, 0}

Thus, for an individual to download, we must have

U = max{ ν – p, α ν + (1 – α)δU – ε, 0} = α ν + (1 – α)δU – ε,

and in order for this to occur, two conditions below need to be satisfied:

1a) u(download) > u(stay idle) or α ν + (1 – α)δU – ε > 0,

2a) u(download) > u(purchase) or α ν + (1 – α)δU – ε > ν – p.

Hence,

$U = \alpha\nu + (1-\alpha)\delta U - \varepsilon \Rightarrow U = \dfrac{(a\nu - \varepsilon)}{(1-\delta(1-\alpha))}$. Inserting this into (1a), we obtain

$$0 < \alpha\nu + \frac{(1-\alpha\nu + (1-\alpha)\delta(a\nu - \varepsilon))}{(1-\delta(1-\alpha))} - \varepsilon$$

$$\frac{\varepsilon}{\alpha} < \nu$$

$$\nu - p < \alpha\nu + \frac{(1-\alpha\nu + (1-\alpha)\delta(a\nu - \varepsilon))}{(1-\delta(1-\alpha))} - \varepsilon$$

$$\nu < \frac{p(1-\delta(1-\alpha)) - \varepsilon}{(1-\alpha)(1-\delta)}$$

After above algebraic manipulations, it can be shown that the above two expressions, 1a), 2a) are equivalent to

$$\frac{\varepsilon}{\alpha} < \nu < \frac{p(1-\delta(1-\alpha)) - \varepsilon}{(1-\alpha)(1-\delta)} \qquad (4)$$

(4) demonstrates that if the individual chooses to download, then, the individual's value of the item, v , must lie between $\frac{\varepsilon}{\alpha}$ and $\frac{p(1-\delta(1-\alpha))-\varepsilon}{(1-\alpha)(1-\delta)}$, and if $\frac{\varepsilon}{\alpha} > \frac{p(1-\delta(1-\alpha))-\varepsilon}{(1-\alpha)(1-\delta)}$, then the above condition automatically fails to hold, so the individual will not download at all.

In reality, the probability depends on the number of uploaders, so I have to modify α, so that it is a function of number of uploaders in equilibrium.

Let's consider the simplest case first, the P2P network of two users: i and j, and assume that they are both looking for a certain item (not necessarily same); whether i or j can download the complete item solely depends on whether the other will upload that item or not. Let's set the price of all the items equal to p and assume that both individuals have same discount factor, δ and bear the same cost, ε, for the sake of simplicity. Let α_i denote the probability that individual i can successfully download the item, and α_j is defined similarly.

Let's define e_i as a probability that an individual i will enter the P2P network, and when the individual enters the network, he or she will automatically participate in downloading. In order to determine the number of uploaders, it is necessary to introduce a new variable, a probability that someone will upload a file that the other wants once he enters the P2P network. γ is defined as a fixed probability that a P2P user will upload the file that the other user wants. This probability may capture similarity for network users' preferences. For instance, if the network is for people who like jazz music, then γ may be higher than the one for people who like music in general.

I assume that all P2P users have the same expectation of "reciprocity," meaning as each P2P user expects that if he uploads an item, it will motivate other people to upload as much as when other P2P user uploads the item, and, therefore, γ is constant in this network. In this case, α_j can be rewritten as

$\alpha_j = \gamma e_i$, and by a similar argument, $\alpha_i = \gamma e_j$.

By combining these modifications with (4), I can deduce that the probability that an individual enters a P2P network is subject to the following constraint,

$$\frac{\varepsilon}{\gamma e_j} < v_i < \frac{p(1-\delta(1-\gamma e_j)) - \varepsilon}{(1-\gamma e_j)(1-\delta)}$$

$$\frac{\varepsilon}{\gamma e_i} < v_j < \frac{p(1-\delta(1-\gamma e_i)) - \varepsilon}{(1-\gamma e_i)(1-\delta)}$$

To keep the model simple, let's assume that ν is uniformly distributed on the interval, [0, $\bar{\mathrm{v}}$] where $\bar{\mathrm{v}}$ is an arbitrary real-valued upper bar greater than zero, and both individuals have the same distribution functions.

In this modified model, α is no longer constant, but a function of e and γ. Let F be a cumulative distribution function such that F(k) calculates a probability that a random variable takes a value less than any real value k greater or equal to 0. Since ν is uniformly distributed in the interval, [0, $\bar{\mathrm{v}}$], F(x) = x/$\bar{\mathrm{v}}$ for ν = x. Mininum F(x) is therefore 0, and since $\bar{\mathrm{v}}$ is the maximum upper bar for ν, F(ν) is never greater than 1. In the model, it's assumed that all possible levels of ν are represented, and, therefore, F(ν) is a monotonically increasing continuous function.

Let the upper limit, $\frac{p(1-\delta(1-\alpha)) - \varepsilon}{(1-\alpha)(1-\delta)}$ be $\bar{\nu}$, and the lower limit, $\frac{\varepsilon}{\alpha}$ be $\underline{\nu}$, then

$e_i = F(\bar{\nu}_i) - F(\underline{\nu}_i) = \dfrac{p(1-\delta(1-\alpha_i))-\varepsilon}{\bar{v}(1-\alpha_i)(1-\delta)} - \dfrac{\varepsilon}{\bar{v}\alpha_i} = \left(\dfrac{p(1-\delta(1-\gamma e_j))-\varepsilon}{(1-\gamma e_j)(1-\delta)} - \dfrac{\varepsilon}{\gamma e_j}\right)\left(\dfrac{1}{\bar{\nu}}\right)$

$e_j = F(\bar{\nu}_j) - F(\underline{\nu}_j) = \dfrac{p(1-\delta(1-\alpha_j))-\varepsilon}{\bar{v}(1-\alpha_j)(1-\delta)} - \dfrac{\varepsilon}{\bar{v}\alpha_j} = \left(\dfrac{p(1-\delta(1-\gamma e_i))-\varepsilon}{(1-\gamma e_i)(1-\delta)} - \dfrac{\varepsilon}{\gamma e_i}\right)\left(\dfrac{1}{\bar{\nu}}\right)$

In equilibrium, $e^* = e_i = e_j$, so given the values of $\bar{\nu}$, p, δ, γ, and ε, we can solve for the probability that the user will download instead of purchasing, using the equation

$e^* = \dfrac{p(1-\delta(1-\alpha^*))-\varepsilon}{\bar{v}(1-\alpha^*)(1-\delta)} - \dfrac{\varepsilon}{\bar{v}\alpha^*}$

or equivalently,

$$e^* = \left(\frac{p(1-\delta(1-\gamma e^*))-\varepsilon}{(1-\gamma e^*)(1-\delta)} - \frac{\varepsilon}{\gamma e^*}\right)\left(\frac{1}{\bar{\nu}}\right) \qquad (5)$$

The above model was the case for N = 2, but most P2P networks have millions of users, so I will extend my model to N > 2 and determine how the equilibrium value of e* depends on N.

The formula, (4), still holds as a necessary condition for an individual to download, but for an individual j, the value of α_j with respect to other parameters changes. When N = 2, it was just the probability that the other person will upload it, but when N > 2, it equals to 1 – Pr(nobody uploads the file), which is same as

$$1-\prod_{x=1}^{N-1}(1-\alpha_x) \text{ or } 1-\prod_{x=1}^{N-1}(1-\lambda e_x) \text{ for } j \neq x \qquad (6)$$

Thus, $\alpha_j = 1-\prod_{x=1}^{N-1}(1-\lambda e_x)$. For simplicity, we focus on symmetric equilibria at which everybody's probability of successfully downloading from the network is equal, and so is the probability of entering P2P network. Let's denote these parameters by the symbol α*, e* respectively, then from (6), we can deduce that $\alpha^* = 1-(1-\gamma e^*)^{N-1}$

Let F(x) be defined similarly as it was in the case of N = 2, and again, ν is uniformly distributed in the interval, [0, $\bar{v}$].

Let the upper limit, $\dfrac{p(1-\delta(1-\alpha^*))-\varepsilon}{(1-\alpha^*)(1-\delta)}$ be $\bar{\nu}$, and the lower limit, $\dfrac{\varepsilon}{\alpha^*}$ be $\underline{\nu}$

Then (5) is equivalent to

$$e^* = \left(\frac{p(1-\delta(1-(1-(1-\gamma e^*)^{N-1})))-\varepsilon}{(1-(1-(1-\gamma e^*)^{N-1}))(1-\delta)} - \frac{\varepsilon}{1-(1-\gamma e^*)^{N-1}}\right)\left(\frac{1}{\bar{\nu}}\right) \quad \text{or}$$

$$e^* = \left(\frac{p(1-\delta(1-\gamma e^*)^{N-1})-\varepsilon}{(1-\gamma e^*)^{N-1}(1-\delta)} - \frac{\varepsilon}{1-(1-\gamma e^*)^{N-1}}\right)\left(\frac{1}{\bar{\nu}}\right) \tag{7}$$

, which is the probability that an individual's value of ν will stay within the constraint (4).

Since F(x) is a probability distribution function, it should never be greater than 1 or smaller than 0, and, therefore, the above expression, (7), only holds if and only if the following two conditions are satisfied.

1b) $\left(\dfrac{p(1-\delta(1-\gamma e^*)^{N-1})-\varepsilon}{(1-\gamma e^*)^{N-1}(1-\delta)} - \dfrac{\varepsilon}{1-(1-\gamma e^*)^{N-1}}\right) \geq 0$ or equivalently,

$\dfrac{p(1-\delta(1-\gamma e^*)^{N-1})-\varepsilon}{(1-\gamma e^*)^{N-1}(1-\delta)} \leq \dfrac{\varepsilon}{1-(1-\gamma e^*)^{N-1}}$, and if not, then $e^* = 0$

2b) $\left(\dfrac{p(1-\delta(1-\gamma e^*)^{N-1})-\varepsilon}{(1-\gamma e^*)^{N-1}(1-\delta)}\right)\left(\dfrac{1}{\bar{\nu}}\right) \leq 1$, and if not, then

$e^* = 1 - \dfrac{\varepsilon}{1-(1-\gamma e^*)^{N-1}}$.

Therefore, (7) holds for only a limited range of N.

Let N_i denote the minimum integer N such that

$\frac{p(1-\delta(1-\gamma e^*)^{N-1})-\varepsilon}{(1-\gamma e^*)^{N-1}(1-\delta)} \geq \frac{\varepsilon}{1-(1-\gamma e^*)^{N-1}}$, and let N_f denote the maximum integer N such

that $(\frac{p(1-\delta(1-\gamma e^*)^{N-1})-\varepsilon}{(1-\gamma e^*)^{N-1}(1-\delta)})(\frac{1}{\nu}) \leq 1$, then (7) holds for $N \in [N_i, N_f]$ for $N \in Z$.

$$e^* = (\frac{p(1-\delta(1-\gamma e^*)^{N-1})-\varepsilon}{(1-\gamma e^*)^{N-1}(1-\delta)} - \frac{\varepsilon}{1-(1-\gamma e^*)^{N-1}})(\frac{1}{\nu})$$

It is important to note that $N_f > N_i$ if and only if the expression (7) is an increasing function of N. Hence, in order to show that $N_f > N_i$, so that the interval is properly defined, I have to show that the derivative of the e* with respect to N is positive.

Let's define $H(e^*, N) = F((\frac{p(1-\delta(1-\gamma e^*)^{N-1})-\varepsilon}{(1-\gamma e^*)^{N-1}(1-\delta)} - \frac{\varepsilon}{1-(1-\gamma e^*)^{N-1}})(\frac{1}{\nu})) - e^*$, then for a fixed value of H,

$$\frac{\partial H}{\partial e}\frac{\partial e}{\partial N} + \frac{\partial H}{\partial N}\frac{\partial N}{\partial N} = 0$$

$$\frac{\partial H}{\partial e}\frac{\partial e}{\partial N} = -\frac{\partial H}{\partial N}\frac{\partial N}{\partial N} = -\frac{\partial H}{\partial N}$$

$$\frac{\partial e}{\partial N} = \frac{-\frac{\partial H}{\partial N}}{\frac{\partial H}{\partial e}}$$

In order to compute the sign of $\frac{\partial H}{\partial N}$, I have to calculate

$$\frac{\partial F((\frac{p(1-\delta(1-\gamma e^*)^{N-1})-\varepsilon}{(1-\gamma e^*)^{N-1}(1-\delta)} - \frac{\varepsilon}{1-(1-\gamma e^*)^{N-1}})(\frac{1}{\nu}) - e^*)}{\partial N} =$$

$$\log(1-\gamma e^*)(\underbrace{\frac{-p}{(1-\delta)}}_{(1^*)} - \underbrace{\frac{(1-\gamma e^*)^{1-N}(p(1-\delta(1-\gamma e^*)^{N-1})-\varepsilon)}{(1-\delta)}}_{(2^*)} - \underbrace{\frac{\varepsilon(1-\gamma e^*)^{N-1}}{(1-\delta)}}_{(3^*)})$$

Since $\log(1-\gamma e^*) < 0$, and we're assuming ε is a very small value,

so $p(1-\delta(1-\gamma e^*)^{N-1}) - \varepsilon > 0$, and since (1*), (2*), (3*) < 0, and, $\frac{\partial H}{\partial N} > 0$. (1')

To calculate the sign of $\frac{\partial H}{\partial e}$, I will analyze how the value of H changes (noted by dH) for an infinitesimal change of e (noted by de)

$$dH = \left(\frac{p(1-\delta(1-\gamma(e+de))^{N-1}) - \varepsilon}{(1-\gamma(e+de))^{N-1}(1-\delta)} - \frac{\varepsilon}{1-(1-\gamma(e+de))^{N-1}}\right)\left(\frac{1}{\nu}\right) - (e+de) -$$

$$\left(\left(\frac{p(1-\delta(1-\gamma e)^{N-1}) - \varepsilon}{(1-\gamma e)^{N-1}(1-\delta)} - \frac{\varepsilon}{1-(1-\gamma e)^{N-1}}\right)\left(\frac{1}{\nu}\right) - e\right)$$

$$= \frac{1}{1-\delta}\Bigg(\underbrace{p((1-\gamma(e+de))^{N-1} - (1-\gamma e)^{N-1})}_{(1^{**})} - \varepsilon\underbrace{\left(\frac{1}{(1-\gamma(e+de))^{N-1}} - \frac{1}{(1-\gamma e)^{N-1}}\right)}_{(2^{**})}$$

$$-\varepsilon\underbrace{\left(\frac{1}{1-(1-\gamma(e+de))^{N-1}} - \frac{1}{1-(1-\gamma e)^{N-1}}\right)}_{(3^{**})} - de\Bigg)$$

For $de > 0$, (1**) is negative, -(2**) is negative, and -(3**) is positive, but since ε is assumed to be sufficiently small, and for a finite range of N, the effects of (2**) and (3**) are neutralized by the effect of (1**), so the overall effect of e on H is negative, and, therefore, $\frac{\partial H}{\partial e} < 0$ (2').

Combining (1') and (2')

$$\frac{\partial e}{\partial N} = \frac{-\frac{\partial H}{\partial N}}{\frac{\partial H}{\partial e}} > 0$$

Therefore, $N_i < N_f$, and $N \in [N_i, N_f]$ is a properly defined interval.

In short, given the values of parameters, N, δ, γ, ε, p, $\bar{\nu}$ I can calculate the probability that a randomly selected individual from a population will enter the network, e*, using the equations

$$e^* = 0 \qquad N \in [0, N_i]$$

$$e^* = \left(\frac{p(1-\delta(1-\gamma e^*)^{N-1}) - \varepsilon}{(1-\gamma e^*)^{N-1}(1-\delta)} - \frac{\varepsilon}{1-(1-\gamma e^*)^{N-1}}\right)\left(\frac{1}{\bar{\nu}}\right) \qquad N \in [N_i, N_f]$$

$$e^* = 1 - \frac{\varepsilon}{1-(1-\gamma e^*)^{N-1}} \qquad N \in [N_f, \infty)$$

In the absence of a legal barrier, the value of ε will be zero, so the above equation becomes

$$e^* = 0 \qquad N = 0$$

$$e^* = \frac{p(1-\delta(1-\gamma e^*)^{N-1})}{(1-\gamma e^*)^{N-1}(1-\delta)\bar{\nu}} \qquad N \in (0, N_{f(absent)}]$$

$$e^* = 1 \qquad N \in [N_{f(absent)}, \infty)$$

In which $N_{f(absent)}$ denotes the maximum integer N such that

$\frac{p(1-\delta(1-\gamma e^*)^{N-1})}{(1-\gamma e^*)^{N-1}(1-\delta)\bar{\nu}} \leq 1$ in the absence of a legal sanction. It will be lower than N since e* will increase faster in the absence of a legal barrier than in the presence of a legal barrier. Therefore, given a non-zero population size N, and after calculating the equilibrium value, e* for both cases (in the presence of a legal sanction and an absence of a legal sanction) a the extent to which a legal barrier will reduce the number of illegal P2P users can be calculated using the methods below

$$N(e^*_{absent} - e^*_{present}) = \frac{Np(1-\delta(1-\gamma e^*_{absent})^{N-1})}{(1-\gamma e^*_{absent})^{N-1}(1-\delta)\bar{\nu}} \qquad \text{for} \qquad N \in (0, N_i]$$

$$N(e^*_{absent} - e^*_{present}) = N((\frac{p(1-\delta(1-\gamma e^*_{absent})^{N-1}) - \varepsilon}{(1-\gamma e^*_{absent})^{N-1}(1-\delta)}) - (\frac{p(1-\delta(1-\gamma e^*_{present})^{N-1}) - \varepsilon}{(1-\gamma e^*_{present})^{N-1}(1-\delta)} - \frac{\varepsilon}{1-(1-\gamma e^*_{present})^{N-1}}))(\frac{1}{\bar{\nu}}) \qquad \text{for } N \in [N_i, N_{f(absent)}]$$

$$N(e^*_{absent} - e^*_{present}) = N(1-(\frac{p(1-\delta(1-\gamma e^*_{present})^{N-1}) - \varepsilon}{(1-\gamma e^*_{present})^{N-1}(1-\delta)} - \frac{\varepsilon}{1-(1-\gamma e^*_{present})^{N-1}})(\frac{1}{\bar{\nu}})) \qquad \text{for } N \in [N_{f(absent)}, N_f]$$

$$N(e^*_{absent} - e^*_{present}) = N(1-(1-\frac{\varepsilon}{\bar{\nu}(1-(1-\gamma e^*_{present})^{N-1}})) = \frac{N\varepsilon}{\bar{\nu}(1-(1-\gamma e^*_{present})^{N-1}} \qquad \text{for} \qquad N \in [N_f, \infty)$$

For a numerical example, let's consider the following case.

A price of a typical music album is about 15 dollars, and let's assume that there exists a very avid fan of a musician, so the value of $\bar{\nu}$ is about 150. Individuals are reasonably patient, so the value for δ is 0.9. The legal fear and the probability that someone will upload what others want are very small values. Thus, the numerical values of the parameters are as follows:

$\bar{\nu} = 150$
$p = 15$
$\delta = 0.9$
$\gamma = 0.01$
$\varepsilon = 0.01$

The tables of e* vs N given those parameters are shown below. Due to the extreme sensitivity of e* to the value of N, it was not possible to find out the value of e* with a high degree of precision. The below charts list the reasonable approximations of values of e* both in the presence of a legal sanction and in the absence of a legal sanction and are subject to an error of ± 0.01. Given the value of N, there are multiple possible values for e*. In order to get a unique solution for e* for each integer value of N, I eliminated the values of e* that do not satisfy $e^* \in [0,1]$, and for N1 < N2, I looked for $\min e^*_{N2}$ such that $e^*_{N1} < e^*_{N2}$. Utilizing this method, I was able to find an approximate value for e* for each integer value of N.

Table 1: e* vs N in the presence of a legal barrier

e*	N
0	<25
0.001	25
0.1	33
0.3	60
0.4	70
0.5	72
0.9	74
$\approx 1 - \frac{0.01}{1-(1-0.01)^{N-1}}$	N>74

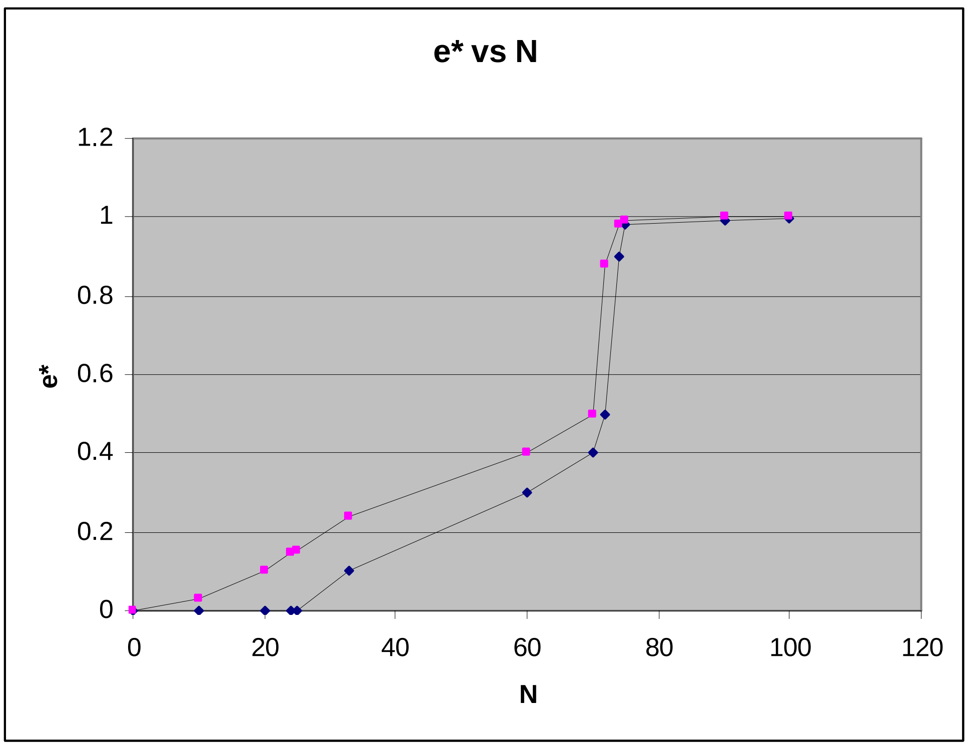
Table 2: e* vs N in the absence of a legal barrier

e*	N
0	< 20
0.1	20
0.3	42
0.4	43
0.5	60
0.9	72
≈ 1	N>72

As shown in the tables, the legal sanction significantly affects e* for very small value of N, but as N gets larger, the effect of the legal sanction pretty much disappears. For instance, at N > 74, the difference between equilibrium value of e* in the presence of the legal barrier and that in the absence of the legal barrier is approximately 0.01.

A sudden increase in e* from 0.5 to 0.9 suggests that once the probability of the successful download exceeds a certain threshold point, individuals enter the network at a rapidly increasing rate.

The graphical representation is shown below, and the pink dots represent the data points in the absence of the legal barrier while the blue dots represent the data points in the presence of the legal barrier.



As you can clearly see, the magnitude of the differential between two values of e* (one in which the legal sanction is enforced, and one where the sanction is absent) decreases as N increases.

Section IV

Conclusion: Even though the probability of successfully downloading the item within the expected period of time reaches 1 as N exceeds a certain number regardless of parameters, some people choose not to download in the presence of a legal sanction. In fact, in the above numerical example, as N approaches infinity, e* approaches

$\lim_{N\to\infty} 1-\frac{0.01}{1-(1-0.01)^{N-1}}$ = 0.99. It is because some individuals value the file very low, so it's not worth for them to risk facing a legal sanction although they can download the file with a 100% certainty.

Overall, the legal barrier, at least in my model, is quite efficient in reducing the P2P users for a very small population size, but when the population size is large enough, it is not effective, and as the value of $\frac{\varepsilon}{\underline{v}(1-(1-\gamma e^*_{present})^{N-1}}$ decreases as N increases, the extent to which the legal barrier prevents individuals from entering the P2P network decreases as well. Since most P2P networks are big, the effect of the legal sanction is minimal overall.

One limitation of the model is that generous uploaders are excluded from the model. If such P2P users are introduced to the network, at a given value of N, the value of α will be greater than that found in the model, so the equilibrium value of e* will be larger, and consequently, the effect of the legal sanction will diminish at each value of N. Since generous uploaders do exist in the real P2P networks, the value of α is greater than that found in the model, and, therefore, the effect of the legal sanction is even smaller.

Another major limitation of the model is that the legal fear is kept constant. In reality, it is actually not. Becker and Clement's paper implies that the legal fear has a high correlation with frequency with which individuals access the P2P network. In fact, the government frequently utilizes the punishment schemes that target specifically at frequent users, making a legal sanction a function of how much individuals upload or download rather than a constant. However, P2P users usually come up with open-source programs that enable them to hide from the government about how much they actually upload or download, and the rate at which such open-source programs are made available to the public is faster than the rate at which government comes up with a new punishment scheme, so if individuals take advantage of those open source programs, the government cannot really pick on frequent users; eventually, this will lead to the government resorting to randomly selecting P2P users. Also, since I'm limiting my model to medium sharers, the amount of the files each sharer uploads or downloads at equilibrium does not vary much from one another, so the variance among the individual legal fears of medium sharers will be very small. Thus, if you're looking at the long-run equilibrium, the legal fear being constant is not an unreasonable abstraction.

Bibliography:

Becker J. and Clement M., (2006) *Dynamics of Illegal Participation in Peer-to-Peer networks, - Why Do People Illegally Share Media Files?*, Journal of Media Economics, 19(1), 7~32

Oberholzer-Gee F. and Strumpf K. (2007) *The Effect of File Sharing on Record Sales *An Empirical Analysis*, Journal of Political Economy 115(1):1-42

Ma R., Lee C.M., Lui J., and Yau D. (2001) *Incentive P2P Networks: A Protocol to Entourage Information Sharing and Contribution*, Columbia University

Ostrom E., Gardner R., & Walker J. (1994) *Rules, Games, and Common pool Resources*, Ann Arbor, University of Michigan Press

Graham J. (2004, July 11) *Online File Sharing Endures*, USA Today