Copying of Information Goods: the Effect of Network Structure on Firm Profits

Mikhael Shor *
Owen Graduate School of Management
Vanderbilt University

Michael R Galbreth †
Moore School of Business
University of South Carolina

September 8, 2008

*mike.shor@owen.vanderbilt.edu, Owen Graduate School of Management, Vanderbilt University, Nashville, Tennessee 37203
†galbreth@moore.sc.edu, Moore School of Business, University of South Carolina, Columbia, South Carolina 29208 (corresponding author)

Copying of Information Goods: The Effect of Network Structure on Firm Profits

Traditionally, illegal copying of information goods has been viewed as detrimental to firms as illicit copies compete directly with new products. In this paper, we present a model of probabilistic group formation for the purpose of copying. The effects of copying on firm profits are examined given two key market characteristics: the network structures that give rise to groups of consumers and the internal cost-sharing mechanisms within groups. We find that illicit copying can increase firm profit when consumers adopt an efficient cost-sharing mechanism. Otherwise, firms prefer the absence of copying; however, when copying activity can only be controlled within a limited range, firms may prefer more copying to less.

1. Introduction

Technological advances are constantly increasing the ease of illicit reproduction and transfer of information property by consumers. Firms generally contend that copyright infringements are detrimental to their profits since goods find their way into the hands of non-paying consumers. Indeed, some have claimed that a piracy rate of just 10% would preclude a recording company from operating profitably (Anand & Galetovic 2004). Industry groups estimate annual losses due to illicit copying at over three billion dollars for DVDs (MPAA 2008), over twelve billion for music (RIAA 2008), and nearly forty billion for software (BSA 2006). These estimates often assume that firms act myopically, setting prices without considering the effect of copying, and facing a profit disappointment when illicit sharing results in reduced sales.

Despite industry claims, researchers are divided on whether illicit copying need reduce firm profits. The literature on copyright infringements by consumers diverges based on whether or not it is assumed that firms are able to price discriminate between individual consumers and groups of consumers who share or copy a product. Where price discrimination is not possible, the argument follows the traditional analysis that individuals who copy a good either have valuations below the current price—in which case they would not become consumers anyway—or have valuations above the firm's profit-maximizing price, hence lowering profit for the firm. Thus, most studies conclude that in the absence of price discrimination, copying can only have deleterious effects on firm profit. For example, Johnson (1985) argues that copying may be welfare improving, but it always reduces firm profit. Johnson assumes that copiers do not purchase originals from which to make these copies, guaranteeing that copying simply decreases the number of potential consumers for the firm. He suggests that the secondary demand effects, which represent the extra amount of revenue generated by a copier's need for an original product from which to make a duplicate, are negligible. However, Johnson's model implies that almost $\frac{1}{3}$ of all consumers copy the product. This may be too sizeable a proportion to justify neglecting secondary demand effects.

Recent literature suggests that secondary demand effects can be important. If firms anticipate sharing among multiple consumers and set prices accordingly, profits may even increase above the level achievable in the absence of copying. Analysis of pricing of scholarly journals to individual subscribers and libraries supports this argument. A firm can increase profit by indirectly appropriating rents from library users through higher prices for libraries than for individual subscribers (Ordover & Willig 1978). Liebowitz (1985) provides some empirical support, demonstrating that the gap between the institutional price and the individual subscription rate broadens with the popularity of the journal.

The context of libraries is somewhat divorced from the general problem of copying by consumers. Most copying is not done in a library environment where monitoring is feasible and it may be fairly simple to exercise effective price discrimination. Copyright holders of music, software, and printed works are affected by the ability of private individuals to share goods without an opportunity for the firm to appropriate any additional revenue. However, there are some similarities between libraries and consumer copying in general. In libraries, raising the price for individual subscriptions leads to fewer purchases by individuals at this higher price, but raises libraries' demand for the journals as former subscribers now use institutions. This inherent tradeoff does not disappear if firms lose the ability to price the two markets separately. We consider the case where consumers form groups in order to purchase a good and share it through copying. This is in contrast to previous studies of pricing for easily copied goods such as Khouja & Smith (2007), Nascimento & VanHonacker (1988), and Sundararajan (2004), which do not address the possibility of group formation to share the cost of a single item and copy it. If copying among consumers cannot be monitored, the firm controls only a single price. However, an increase in that price, while lowering demand from individuals who are not part of groups, can capture greater surplus from collections of individuals who do copy.

In this manuscript, we investigate copying's impact on firm profits, explicitly considering two factors—the social networks through which copying groups probabilistically form, and the methods by which groups make collective purchasing and price-sharing decisions.

2. Social Networks

As noted by Backstrom et al. (2006), the groups that form within an underlying network of individuals can be thought of as subgraphs of that network, the growth of which can be quite complex and difficult to capture empirically. Groups that facilitate sharing and copying, as with other networks describing interpersonal relations, may take on several representations. Baker & Faulkner (1993), in an empirical study of the organization of groups formed for illicit activities, conclude that the need to conceal these activities generally favors a decentralized network, but high information-processing needs required for some illegal arrangements may result in the formation of more centralized networks. Since copying is often an illicit activity, either of these network approaches may be justified depending on the relative importance of concealment and information needs. Generally, networks established for illegal activities can be quite diverse in their topology (Fijnaut et al. 1998).

Several studies of illicit copying note the importance of social network structure. Limayem et al. (2004) find that, in the case of software piracy, access to the software to be copied and knowing someone who can help with pirating were both significant factors explaining illegal software copying. This implies that group formation is an important aspect of illicit copying of information goods. The idea of group formation for the purpose of copying information goods was established in the literature by Gopal & Sanders (1998), but that work assumed that consumers would deterministically join into groups of optimal size given the pricing and control decisions of the firm. In contrast, our analysis allows for probabilistic group formation and focuses on the resulting distribution of group sizes.

In the economics literature, prior research has shown that the formation of consumer groups for the purpose of sharing can lead to higher profits for firms, even in the absence of price discrimination (Bakos et al. 1999, Varian 2005). Bakos et al. (1999) find that firms can profit from consumer sharing because sharing "reshapes demand" by pooling consumer valuations into groups. Since each group is essentially a single consumer, the variance in the distribution of consumer valuations faced by the firm is reduced and profits can be increased. However, the increased variation caused by heterogeneity in group sizes can lead to decreased profits if it overwhelms the effects of aggregating consumers into groups. Thus, the critical consideration in determining the impact of piracy is the topology of consumer groups.

We add to the understanding of copying's impact on firm profits by employing graph theory to determine group size distributions. The representation of complex social networks as graphs dates back over fifty years (Cartwright & Harary 1956). Individuals are represented as nodes and the existence of an edge connecting two individuals represents a potential relationship. A random graph (Bollobas 2001) assigns a probability to each potential relationship. In our context, the probability represents the likelihood that two people both know each other (perhaps virtually) and would share an information good acquired by one of them. Additional considerations that might be relevant to group formation in other settings, such as contingent transfers or payments to prevent some edges from forming (Bloch & jackson 2007) are outside the scope of this study.

We consider three categories of social networks: decentralized and centralized (as suggested by Baker & Faulkner 1993), as well as a complete network. A decentralized network may represent informal copying by word of mouth (e.g., sharing movies or software with friends) in which no central agency provides the copied goods. A centralized network is more formal, perhaps operated for profit or within large organizations. Finally, a complete network represents the case where every consumer has an equal chance of being directly connected to every other consumer.

To highlight the importance of group formation on firm profits, consider the simple example presented in Table 1. A profit-maximizing firm faces four consumers for an information good with valuations of \$8, \$11, \$12, and \$16. In the absence of copying, the firm may sell to all consumers at a price of \$8, to three at a price of \$11, to two at a price of \$12, or to a single

Table 1: Copying Groups and Firm Profit

	Network Structure	Optimal Price	Resulting Quantity	Firm Profit	Consumer Surplus	Effect of Copying
No copying						
	$8, $11, $12, $16	$11	3	$33	$6	
Copying						
i.	($8, $11), ($12, $16)	$19	2	$38	$9	$\Pi \uparrow$ CS $\uparrow$
ii.	($8, $16), ($11, $12)	$23	2	$46	$1	$\Pi \uparrow$ CS $\downarrow$
iii.	($12, $16), $8, $11	$28	1	$28	$0	$\Pi \downarrow$ CS $\downarrow$
iv.	($8, $11, $12), $16	$16	2	$32	$15	$\Pi \downarrow$ CS $\uparrow$

consumer at a price of $16. The firm (which, for simplicity, we assume produces with zero marginal costs) would set a price of $11, earning a profit of $33, and resulting in $6 of consumer surplus. Now consider consumers forming groups, with group members costlessly and without any loss of quality duplicating and sharing the good. Alternate mechanisms for sharing the cost of the good are described in the next section. In this example, we assume for simplicity that a group will purchase the good whenever the sum of members' valuations exceeds the price. Consider the first partition of consumers, (i.), with two groups, ($8,$11) and ($12,$16). To fix ideas, imagine the first group consists of two brothers, and the second of two sisters unrelated to the first group. Each set of siblings can share the item by making an illicit copy, but neither shares with members of the other group because they are unacquainted. Since the total valuations of the two groups are $19 and $28, the firm sets a price of $19, selling to both groups, and earning a profit of $38. Resulting consumer surplus is $9. Both the firm and consumers are benefited by the ability to copy. If the two brothers are unable to copy (perhaps due to physical separation or a sibling dispute), the firm faces the partition in (iii.). The firm would set a price of $28, selling a single unit. Both firm profit and consumer surplus decrease as a result of consumers' ability to copy, implying that the ability to copy need not benefit consumers. The group configurations in Table 1 show that all combinations of increasing and decreasing profits and increasing and decreasing consumer surplus are possible due to the introduction of copying.

Realistically, a firm is unlikely to know which group configuration will occur. Instead, it must consider the likelihood that each of the above groups, and others, will form. Since some group profiles lead to increased profits from copying and others lead to a deterioration in profits, the structure and relative probability of different groups are likely to be important.

3. Group Mechanisms

Once a group of consumers forms, it must make two related decisions: whether or not to purchase the product, and how to divide the costs. If the group consists of friends who willingly share their private valuations for the good, both decisions are relatively simple; the group should buy the good whenever the sum of the members' values exceeds the price, and share the costs in such a way that no one is paying more than his or her value. When there are no norms or relationships that motivate the revelation of true valuations, a classic elicitation problem arises. The group may seek a mechanism to query its members as to their preferences, raise funds for purchase, and divide the price based on the reported values of each member. Since each member's share of the costs may depend on the value that she claims to have for the good, she has an incentive to misrepresent her value. The mechanism the group uses to determine the sharing of the cost must overcome these incentive issues.

A classic result in public goods economics holds that no mechanism exists which is simultaneously efficient (leading to purchase whenever the sum of the values exceeds the price) and budget balancing, requiring payments which, in total, exactly equal the price (Green & Laffont 1977). Thus, authors have concentrated either on mechanisms that are efficient, but may raise too much or too little money (Clarke 1971, Groves 1973, Groves & Ledyard 1977), or mechanisms that are budget balancing, but may lead to foregone purchases even though group members' values, in total, exceed the price (Moulin 1994, Serizawa 1999). We include both efficient and budget balancing approaches in our analysis.

First, consider either a group whose members choose to cooperate and reveal their true valuations for the product or who devise a mechanism which leads to truthful revelation in equilibrium. While many methods exist for dividing the cost among group members (e.g., the Shapley value for cooperative groups or Groves-Ledyard for noncooperative ones), for our purposes it is sufficient to note that these mechanisms are *efficient*— a purchase occurs precisely when the sum of all individual valuations exceeds the price.

Second, a noncooperative group may choose a mechanism which guarantees *budget balancing*, but may lead to an (inefficient) choice not to purchase despite the total group value exceeding the price. Finding an optimal mechanism from among those that balance the budget is the subject of active research. One class of mechanisms in particular has been shown to have strong efficiency properties and uniquely satisfy several desirable conditions (Dearden & Einolf 2003, Norman 2004).[1] The mechanism is also intuitively simple and quite easy to implement, involving each willing group member paying an equal share of the price. Effectively, a group of size n first asks if each member is willing to pay p/n, where p is the price. If all agree, the good is purchased and each pays an equal share. If there is not universal agreement, the group checks if at least $n-1$ people are willing to share the good and each pay $p/(n-1)$, and so on. While this mechanism is clearly budget balancing, it does not always lead to efficient decisions. For example, a group of two consumers with values of \$4 and \$8 would not purchase a good priced at \$10 despite their total values exceeding the price. From the perspective of the firm, this group does not fully aggregate members' values, and thus is less profitable than a group that does.

In the next section, we present models of both efficient and budget balancing mechanisms, as well as the three group topologies described in the previous section. This allows us to investigate a wide range of possible copying environments and their impact on firm profits.

4. Model

4.1 Graphs of Social Networks

A random graph $\Gamma = \{\mathbf{N}, \mathbf{E}, q\}$, prescribes a set of consumers, $\mathbf{N} = \{1, \ldots, N\}$, and a set of edges, $\mathbf{E}$, representing possible linkages between consumers, each occurring independently with probability q. A network structure, h, is a set of the sizes of connected components in a realization of Γ, representing the partition of consumers into groups. Therefore, h is an integer partition of N. For example, $h = \{3, 3, 2\}$ is an economy of eight consumers partitioned

[1] The conditions required are (i) consumers have incentive to report honestly, (ii) consumers choose to participate voluntarily, (iii) "non-bossiness" which rules out consumers having the ability to manipulate others' cost shares without changing their own, and (iv) anonymity, implying that the mechanism treats consumers with identical values equally.

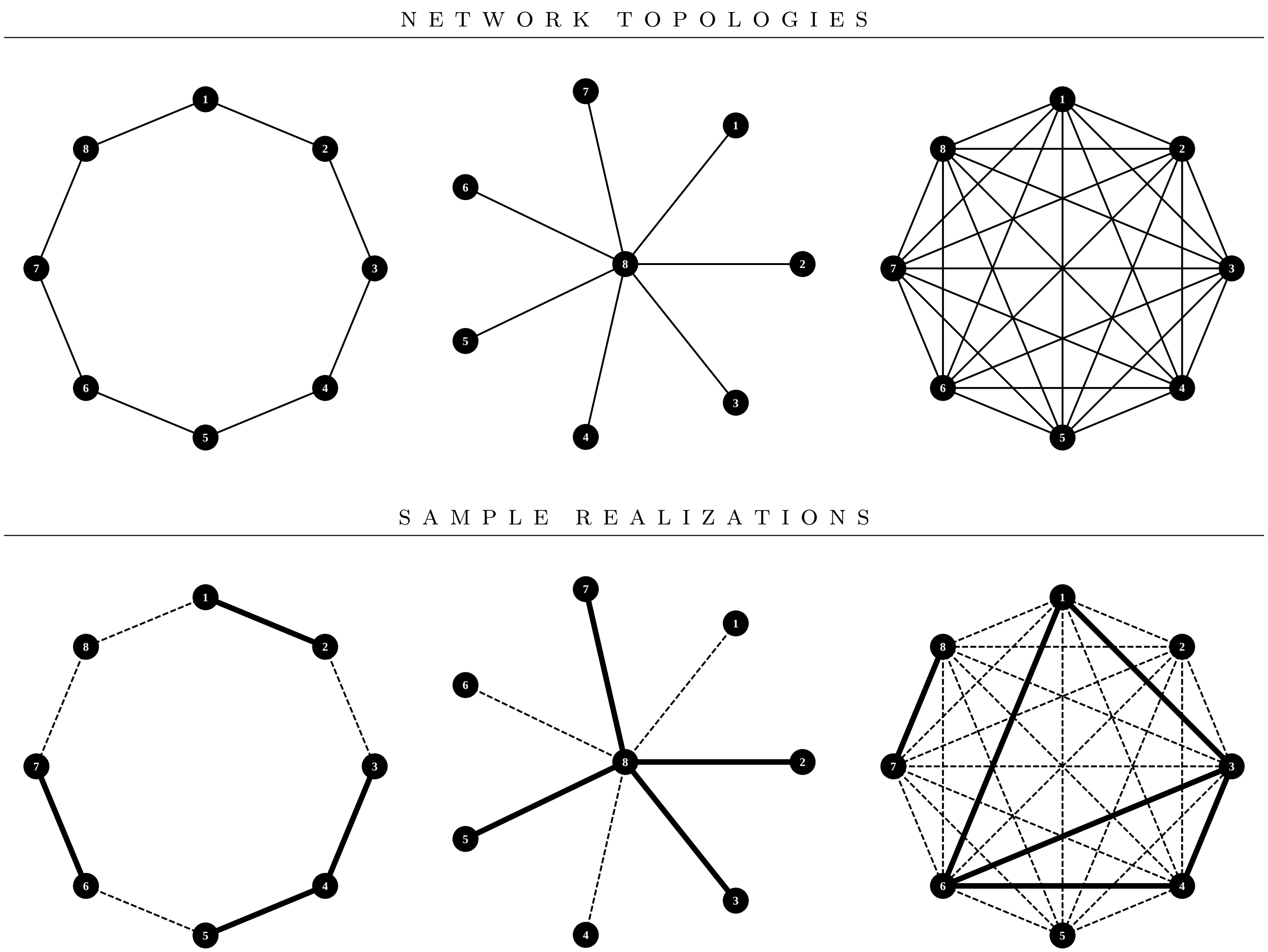
into three groups—two with three members, and one with two members. We consider three topologies of random graphs, represented by the sets of vertices:

- Decentralized Graph: $\mathbf{E} = \{(1,2),(2,3),\ldots(N-1,N),(N,1)\}$
- Centralized Graph: $\mathbf{E} = \{(1,N),(2,N),\ldots(N-1,N)\}$
- Complete Graph: $\mathbf{E} = \{(1,2),\ldots,(1,N),(2,3),\ldots(2,N),\ldots,(N-1,N)\}$

The decentralized network is modeled as a "ring topology," in which a group is formed by a continuous connection of neighboring vertices. Baker & Faulkner (1993) note that those engaged in illicit activities have a tendency to create "sparse and decentralized networks." The ring topology restricts any person's contact to at most two other individuals, limiting the exposure of group members to law enforcement if any one is caught. Figure 1(a) provides an example for an economy of eight consumers. In the sample realization, represented by darkened edges, the firm faces four groups: two aggregating the values of two consumers, one consisting of three consumers, and one singleton.

In contrast to the decentralized graph, a centralized graph (or "star topology") represents a single central entity with whom all other agents are connected. Figure 1(b) presents an example. Since all peripheral nodes may only share through the central node, the star topology represents the most centralized social network. Under several models of endogenous network formation, the star topology is stable, in the sense that no participants have incentives to create or destroy any links (e.g., Jackson & Wolinsky 1996) and may be socially optimal when connections are costly (Fabrikant et al. 2003). Technologically, it represents a client-server framework, such as a network computer devoted to hosting illegal copies of music. The person at node N is the center, and is connected (probabilistically) to each of the remaining consumers, none of whom have any other links. Note that this topology gives rise to at most one group with multiple members.

Lastly, a complete graph, pictured in Figure 1(c), has each consumer connected to every other consumer with equal probability. A firm with no prior knowledge about consumers' network structure may reasonably assume that all consumer linkages are equally likely, giving rise to this graph structure.



RESULTING NETWORKS

{1,2}, {3,4,5}, {6,7}, {8}	{2,3,6,7,8}, {1},{4},{5}	{1,3,4,6}, {7,8},{2},{5}
$h = \{3, 2, 2, 1\}$	$h = \{5, 1, 1, 1\}$	$h = \{4, 2, 1, 1\}$
(a) Decentralized Graph	(b) Centralized Graph	(c) Complete Graph

Figure 1: Network topologies, with example realizations

4.2 Group Mechanisms

Each consumer has a valuation independently distributed according to the distribution function, $F(v)$. Groups adopt one of two mechanisms for deciding whether to purchase and determining members' cost shares: either an efficient or a budget balancing mechanism. If groups adopt an efficient mechanism, then they purchase whenever the sum of their values exceeds the purchase price. From the firm's standpoint, a group of size k has valuation $F_k(v) = v_1 * \ldots * v_k$, which is the k-fold convolution of F, representing the sum of the members' values. If groups adopt the budget balancing mechanism, then a purchase is made whenever there exist m group members who are willing to pay at least p/m. Denote by $v_{(i,k)}$ the i^{th} largest realized value in a group of size k. Then, the maximum a group of size k would pay is equal to

$$F_k(v) = \max_{m \leq k} m v_{(m,k)}.$$

4.3 Firm Profits

Consider a monopolist who produces a product at zero marginal cost. In the absence of copying, a firm maximizes $\Pi(p) = p(1 - F(p))$, the price times the chance of purchase. In the presence of copying, the firm's problem is substantially more complex. A network structure yields a probability distribution over possible sizes of groups, and the mechanism employed by group members gives rise to a probability distribution over the maximum willingness to pay for each group. This appears to create a computational difficulty, since the resulting probability distribution is a random convex mixture of convolutions. However, as the following proposition demonstrates, the profit function may be significantly simplified. Network structure enters the profit function only through the expected number of groups of each size.

Proposition 1. *Denote by $F_k(v)$ the distribution function of the valuation of a group of size k and denote by τ_k the expected number of groups of size k. Then,*

$$\Pi(p) = p \sum_{k=1}^{N} (1 - F_k(p))\, \tau_k \tag{1}$$

The result suggests that the quantity demanded is the sum of the expected number of groups of a given size times the probability that a group of that size purchases the good. Firm profits depend on the graph topology only to the extent that different topologies give rise to different τ_k. This allows us to represent copying networks in a fairly straightforward manner, requiring only a set of distributions of group values and a vector representing the expected number of groups of each size.

The chance of any possible link actually occurring is given by q. Clearly, when $q = 0$, no illicit copying occurs as no groups form. When $q = 1$, the resulting graph is connected under all network structures, implying that consumers will purchase at most one copy which will then be shared among the consumers. Naturally, $\sum_{k=1}^{N} k\tau_k = N$. When $q = 0$, $\tau_1 = N$ and when $q = 1$, $\tau_N = 1$. The following proposition derives τ_k for the decentralized and centralized network structure.

Proposition 2. *For the decentralized graph:*

$$\tau_k = Nq^{k-1}(1-q)^2, k \in \{1, \ldots, N-1\}$$
$$\tau_N = Nq^{N-1}(1-q) + q^N$$

For the centralized graph:

$$\tau_k = \binom{N-1}{k-1} q^{k-1}(1-q)^{N-k}, k \in \{2, \ldots, N\}$$
$$\tau_1 = (N-1)(1-q) + (1-q)^{N-1}$$

Explicit solutions for τ_k cannot be derived in closed form for the complete graph,[2] but we can obtain them numerically. For large networks, the complete graph behaves similarly to our centralized network for all but very small q, but attains a high level of connectivity.[3] Lastly, the firm's profit-maximizing price, $p^* \in \text{argmax}\Pi(p)$, obtained by differentiating (1), is given

[2]Finding a complete subgraph of size k was one of the first problems shown to be NP-complete (Karp 1972). In particular, solutions involve transcendental equations, a problem that plagues many even simpler models of random graphs (Newman 2003).

[3]Asymptotically, τ_k has a discontinuity at $q = 1/n$. For $q < 1/n$, the graph contains only small connected components, and for $q > 1/n$, the graph has precisely one large component with all other components small. When $q > \frac{\ln n}{n}$, the graph is almost surely connected (Erdös & Rényi 1960).

implicitly by:

$$p^* = \sum_{k=1}^{N} (1 - F_k(p^*))\, \tau_k \bigg/ \sum_{k=1}^{N} f_k(p^*)\tau_k$$

5. Results

As copying becomes more likely (as q increases), groups of larger sizes form with greater probability, increasing the expected number of larger groups. Simultaneously, an increase in copying activity implies a decrease in the number of groups in the market. If groups can be likened to households, firms face fewer potential households as copying increases but the amount of residents in each household, and thus each household's value for the good, increases. Evaluating the change in network structure as q increases, we have:

$$\begin{array}{lll} \frac{d\tau_1}{dq} < 0, \qquad \frac{d\tau_N}{dq} > 0 & & \text{for all networks} \\ \frac{d\tau_k}{dq} > 0 \Leftrightarrow k > \frac{1+q}{1-q}, & k \in \{2, \ldots, N-1\} & \text{for the decentralized network} \\ \frac{d\tau_k}{dq} > 0 \Leftrightarrow k > (N-1)q + 1, & k \in \{2, \ldots, N-1\} & \text{for the centralized network} \end{array}$$

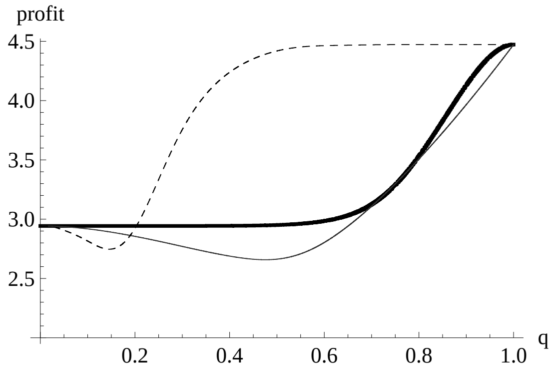
To derive specific results on profit, we first must specify a distribution for consumer valuations. We adopt the exponential distribution for v.[4] Due to the non-transcendental nature of this and similar functions involving convolutions of random variables bound from below, an explicit closed-form solution for the optimal price is not possible.[5] We calculate optimal profits numerically.

Profit curves as a function of q with efficient cost sharing are presented in Figure 2. For decentralized social networks, the optimal profit of the firm is everywhere increasing in q.

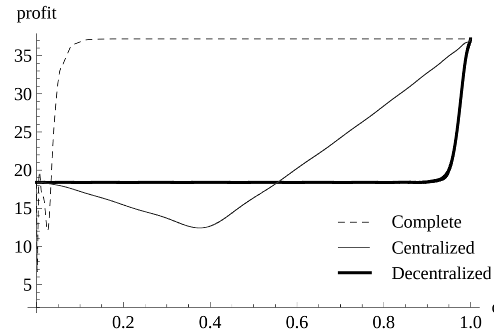
[4]Similar results as those reported were obtained numerically for several parameterizations of the gamma, chi-square, and Weibull distributions. However, the exponential distribution simplifies the calculation of convolutions since the sum of exponentials is given by the gamma distribution. Specifically, denoting by f_k the density of the k-fold convolution,

$$f_k(v) = \frac{v^{k-1}}{(k-1)!} e^{-v} \text{ and } F_k(v) = 1 - \sum_{i=1}^{k} \frac{v^{i-1}}{(i-1)!} e^{-v} = 1 - \sum_{i=1}^{k} f_i(v)$$

[5]This problem can be avoided and explicit results obtained if we consider a distribution on an unbounded domain. However, allowing for negative valuations implies that the addition of another consumer to a group can lead to a decrease in the group's overall valuation. This is not a reasonable reflection of reality.



(a) $N = 8$



(b) $N = 50$

Figure 2: Profit functions for efficient cost sharing under different network structures.

As copying becomes more likely, profits rise, suggesting that the marginal increase in group valuations has a greater impact than the decrease in the number of groups. To counteract the decrease in the number of groups as q rises, the firm raises its price. Individuals not part of a group are less likely to purchase, but the higher price extracts more of the surplus from consumers in groups, leading to increased profits. Figure 2 suggests that profits are little changed in the decentralized case for low values of q (though they are increasing). This can be attributed to the tradeoff between raising prices to capture added surplus of large groups and maintaining prices to not lose customers who have not yet joined larger groups. For low copying activity, these effects appear to balance out in decentralized networks. As copying activity rises, the firm becomes relatively certain that most consumers are members of non-trivial groups, allowing prices to rise accordingly and capture more of the surplus.

For the centralized and complete social networks in Figure 2, the effects are less clear cut. First, consider the case of a centralized graph. Low levels of copying activity lead to decreases in firm profit compared to no copying. This is in agreement with the literature on institutional journal subscriptions. The central agent, the single person through whom all other consumers may be connected, serves as a de facto library—the sole source of a good which consumers may then duplicate. Low levels of copying correspond to some consumers joining this central agent's group. In the absence of price discrimination between individuals and groups, this leads to a reduction in revenues. As the number of individuals who join the group rises, the firm may begin pricing as if it sold only to groups. Hence, there is a critical point of copying at which

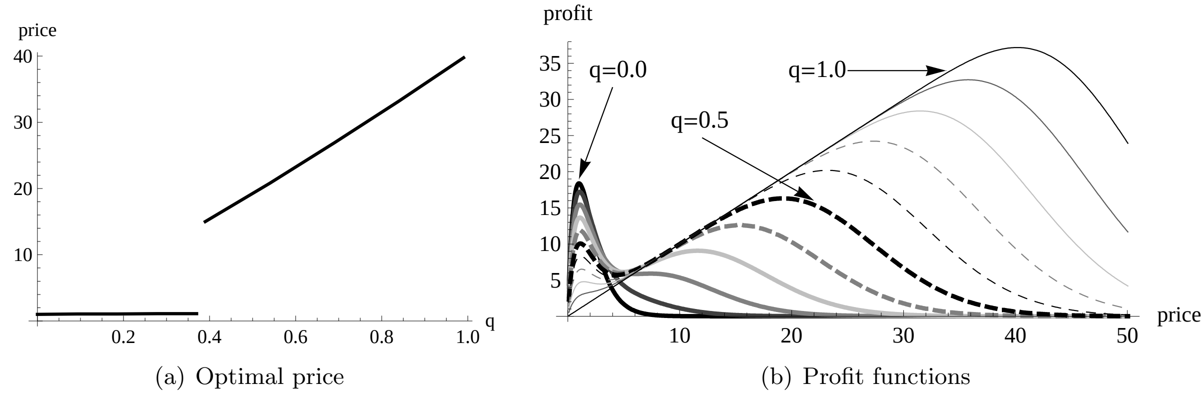


Figure 3: Optimal prices and profit curves for a centralized network with efficient cost sharing as a function of copying activity ($N = 50$).

pricing to maximize profits over individual purchasers ceases to be optimal, and the firm begins charging what are, in effect, institutional rates.

The tradeoff between pricing for individuals or groups is demonstrated in Figure 3(a) which shows a discontinuity in optimal prices. For low q, the firm targets prices toward individuals, and then prices jump upward at a critical value of q. This jump comes about from the double-peaked nature of firm profit as a function of price, as Figure 3(b) illustrates. Profit functions for extreme values of q ($q = 0$ or 1) are well-behaved and single-peaked. Profit functions at other values of q have two peaks, representing the optimal profits from charging individual rates and from charging library or institutional rates. As q rises, the first peak becomes lower, as fewer unconnected consumers remain, and the institutional peak becomes higher, representing greater library usage. The firm must select one of the two peaks. At some critical value of q, institutional selling becomes more profitable.

For the complete graph with an efficient cost sharing mechanism, we see a profit impact similar to the centralized case—profits decrease with copying initially, until a critical point is reached and copying activity becomes profit-increasing. However, with a complete graph, these dynamics are much more concentrated in the low values of q, with profits quickly rising to nearly maximum levels for relatively modest values of q.

When a budget balancing group mechanism is used, groups do not fully appropriate their members' values, and copying is less beneficial for the firm. As shown in Figure 4, the profit-maximizing amount of copying is zero in the budget balancing case regardless of network type.

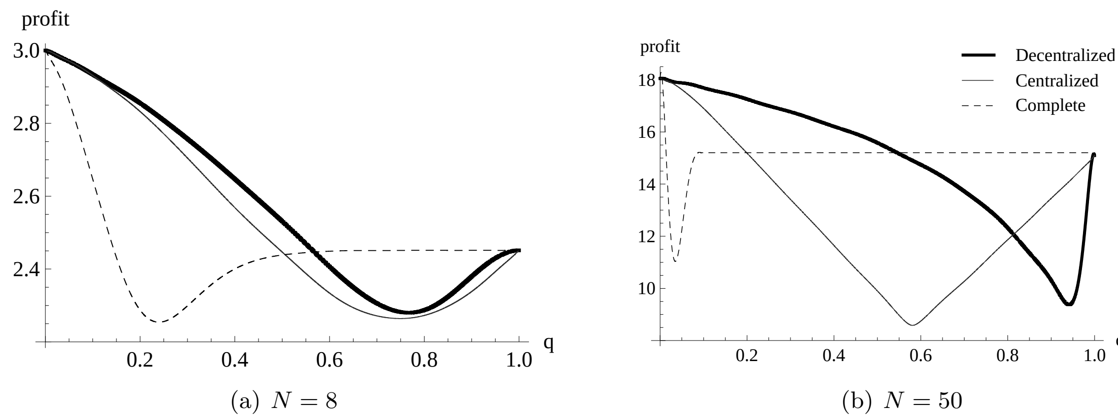


Figure 4: Profit functions for budget balancing cost sharing under different network structures.

Moreover, the negative consequences of copying can be very substantial, with certain levels of copying activity decreasing profits by more than 50%. However, as with the efficient mechanism, there are ranges of q in each network type over which an increase in copying can raise firm profits.

6. Discussion

When cost sharing within copying groups is budget balancing, our results are consistent with the conventional wisdom that profit-maximizing firms prefer a complete absence of copying. However, when costs of an information good are shared efficiently within a group, we find that the complete elimination of copying activity is *never* optimal for a firm. In fact, any increase in copying always increases firm profits if the sharing network is decentralized in nature, e.g. friends copying a music CD. For other network types with efficient cost sharing, profits decrease when copying activity is low, but higher levels of copying allow the firm to adopt higher, institutional prices, appropriating an even greater surplus than in the absence of copying.

For all combinations of network type and sharing mechanism, we find that profits increase over certain ranges of copying activity. Thus, the impact of a *change* in the level of copying is never as simple as "less is better," but rather it depends on current copying levels, which firms may have limited ability to influence. Large swings in q are not likely to be within a company's control; both extremes of complete elimination of illicit copying and nearly a 100% probability that all consumers will participate in copying groups are infeasible. Given this reality, consider profits from the centralized network structure in Figure 4. If q is currently 0.45, and the firm estimates that this can be altered only within a limited range (say plus or minus 5 percentage points), then the firm prefers to decrease copying probability. Enforcement may be justified in these cases, depending on its costs. On the other hand, if q is already high, perhaps at 0.85, then decreased copying would actually be detrimental to firm profits. Although zero copying is globally optimal in this environment, a slight increase in q would enable more profits than a slight decrease.

In practice, firms might find the estimation of q to be difficult. In our model, we can calibrate q to data on the number of illegal copies in a particular industry.[6] For example, one industry group (BSA 2006) estimates that 35% of all software on personal computers is pirated. For several models, we derive the q that gives rise to this proportion in expectation:

- Decentralized, Efficient: $q = 0.214$
- Decentralized, Budget-Balancing: $q = 0.233$
- Centralized, Efficient: $q = 0.168$
- Centralized, Budget-Balancing: $q = 0.170$

We can draw two implications from the numbers above. Locally, all except the efficient decentralized model imply that profits are decreasing in q in the neighborhood of these values. Therefore, at present values, efforts to combat piracy are likely justified. However, the commonly-heard arguments that illicit copying will lead firms to charge substantially higher prices on legal copies (to "recoup" the losses) is suspect. In both centralized models, these parameters lie in the region where increased copying has very little impact on the optimal price.Since losses cannot be recouped by adjusting prices of legal copies at the current level of copying activity, a price increase is not an optimal response in the software industry. In fact, our model suggests a way to uncover the impact of increased copying by observing a firm's optimal price response. If a copying network is best captured by the centralized graph, then when a firm reacts to increased

[6]The expected proportion of illegal copies is given by $1 - \sum \tau_k(1 - F_k(p))/\sum k\tau_k(1 - F_k(p))$, evaluated at $p^*(q)$. It can be confirmed analytically that this measure is increasing in q for the two network structures derived in Proposition 2. Under no copying, $\tau_1 = N$ and the above measure is zero; when $\tau_N = 1$, the measure reduces to $(N - 1)/N$. For the calibration, we use $N = 50$.

copying activity by significantly increasing prices, our model suggests that it is in the region where copying increases firm profits. At this point, the firm has no incentive to marginally decrease copying activity, since it has in effect moved to an institutional pricing model, and prices and profits will continue to increase with more copying. On the other hand, if a firm's reaction to increased copying is to leave prices largely unchanged, then it is likely in a region where profitability is hurt by the additional copying activity, and the incentive to combat illicit copying remains. In short, price increases in the presence of copying signal a firm profiting from, not adversely impacted by, increased copying activity.

Our results indicate that strong action to combat illicit copying, while valid in many contexts, is not always justified. This is especially timely given that enforcement of copyright is becoming increasingly challenging, particularly in some international markets where government support is minimal (Gopal & Sanders 1998). Even where enforcement is possible, it has been noted that piracy prosecution presents challenges given that "suing one's own customers is a highly problematic business strategy" (Maltz & Chiappetta 2002). We find that copyright enforcement is not always profit-enhancing, and perhaps should be avoided in some instances. If consumers form groups in a decentralized manner and share costs efficiently, it is never profit-increasing to lower the probability of copying. In other contexts, the profit impact of reducing copying is mixed, depending both on the current copying activity and the magnitude of the change that enforcement activities would produce.

Appendix

Proof of Proposition 1. Let $H(N)$ be the set of all integer partitions of N with arbitrary element $h_i = \{n_i^1, \ldots, n_i^{s_i}\}$. Denote the maximum willingness to pay of a group by $v_{n_i^j}$.

$$\begin{aligned}
\Pi(p) &= p \sum_{i=1}^{|H|} \Pr\{h_i\} \sum_{j=1}^{s_i} \Pr\{p \leq v_{n_i^j}\} \\
&= p \sum_{i=1}^{|H|} \Pr\{h_i\} \sum_{k=1}^{N} \Pr\{p \leq v_k\} \left|\{j|n_i^j = k\}\right| \\
&= p \sum_{k=1}^{N} \Pr\{p \leq v_k\} \sum_{i=1}^{|H|} \Pr\{h_i\} \left|\{j|n_i^j = k\}\right| \\
&= p \sum_{k=1}^{N} (1 - F_k(p))\tau_k
\end{aligned}$$

□

Proof of Proposition 2. *(i) centralized graph:* note that at most one connected component of size $k > 1$ may exist, and its probability is precisely the chance that $k-1$ of $N-1$ edges are realized. For τ_1:

$$\begin{aligned}
\tau_1 &= N(1-q)^{N-1} + \sum_{k=2}^{N-1} (N-k)\tau_k \\
&= N(1-q)^{N-1} + (N-1) \sum_{k=2}^{N-1} \tau_k - \sum_{k=2}^{N-1} (k-1)\tau_k \\
&= N(1-q)^{N-1} + (N-1)\left(1 - (1-q)^{N-1} - q^{N-1} - q + q^{N-1}\right) \\
&= (N-1)(1-q) + (1-q)^{N-1}
\end{aligned}$$

(ii) decentralized graph: A network of size N is obtained if $N-1$ or N edges are realized. For $k < N$:

$$\begin{aligned}
\tau_k &= \sum_{s=k-1}^{N-2} \frac{N(N-k-1)! q^s (1-q)^{N-s}}{(s-k+1)!(N-s-2)!} \\
&= \sum_{j=0}^{N-k-1} \frac{N(N-k-1)! q^{j+k-1} (1-q)^{N-j-k+1}}{j!(N-j-k-1)!} \\
&= N q^{k-1} (1-q)^{N-k+1} \sum_{j=0}^{N-k-1} \binom{N-k-1}{j} \left(\frac{q}{1-q}\right)^j \\
&= N q^{k-1} (1-q)^2
\end{aligned}$$

□

References

Anand, B. & Galetovic, A. (2004), 'How market smarts can protect property rights', *Harvard Business Review* **82**(12), 72–79.

Backstrom, L., Huttenlocher, D., Kleinberg, J. & Lan, X. (2006), 'Group formation in large social networks: Membership, growth, and evolution', *Proceedings of the 12th ACM SIGKDD Intl. Conf. on Knowledge Discovery and Data Mining* .

Baker, W. E. & Faulkner, R. R. (1993), 'The social organization of conspiracy: Illegal networks in the heavy electrical equipment industry', *American Sociological Review* **58**(6), 837–860.

Bakos, Y., Brynjolfsson, E. & Lichtman, D. (1999), 'Shared information goods', *Journal of Law and Economics* **42**(1), 117–155.

Bloch, F. & jackson, M. O. (2007), 'The formation of networks with transfers among players', *Journal of Economic Theory* **133**(1), 83–110.

Bollobas, B. (2001), *Random Graphs*, Cambridge Mathematical Library, Cambridge University Press.

BSA (2006), Fourth annual bsa and idc global software piracy study. `http://www.bsa.org/globalstudy`. Accessed 02/21/2008.

Cartwright, D. & Harary, F. (1956), 'Structural balance: A generalization of heider's theory', *Psychological Review* **63**(5), 277–293.

Clarke, E. H. (1971), 'Multipart pricing of public goods', *Public Choice* **11**, 17–33.

Dearden, J. & Einolf, K. (2003), 'Strategy-proof allocation of fixed costs', *Review of Economic Design* **8**(July), 185–204.

Erdös, P. & Rényi, A. (1960), On the evolution of random graphs, *in* 'Publication 5. Institute of Mathematics', Hungary: Hungarian Academy of Sciences, pp. 17–61.

Fabrikant, A., Luthra, A., Maneva, E., Papadimitriou, C. H. & Shenker, S. (2003), 'On a network creation game', *Proceedings of 2003 PODC* pp. 347–351.

Fijnaut, C., Bovenkerk, F., Bruinsma, G. & van de Bunt, H. (1998), *Organized Crime in the Netherlands*, The Hague: Kluwer Law International.

Gopal, R. D. & Sanders, G. L. (1998), 'Software piracy: Key issues and impacts', *Information Systems Research* **9**(4), 380–397.

Green, J. & Laffont, J. (1977), 'Characterization of satisfactory mechanisms for the revelation of the preferences for public goods', *Econometrica* **45**(2), 427–438.

Groves, T. (1973), 'Incentives in teams', *Econometrica* **41**, 617–631.

Groves, T. & Ledyard, J. (1977), 'Optimal allocation of public goods: A solution to the 'free-rider' problem', *Econometrica* **45**(4), 783–809.

Jackson, M. O. & Wolinsky, A. (1996), 'A strategic model of social and economic networks', *Journal of Economic Theory* **71**(1), 44–74.

Johnson, W. R. (1985), 'The economics of copying', *The Journal of Political Economy* **93**(1), 158–174.

Karp, R. M. (1972), Reducibility among combinatorial problems, *in* R. E. Miller & J. W. Thatcher, eds, 'Complexity of Computer Computations', New York: Plenum, pp. 85–103.

Khouja, M. & Smith, M. A. (2007), 'Optimal pricing for information goods with piracy and saturation effect', *European Journal of Operational Research* **176**(1), 482–497.

Liebowitz, S. J. (1985), 'Copying and indirect appropriability: Photocopying of journals', *The Journal of Political Economy* **93**(5), 945–957.

Limayem, M., Khalifa, M. & Chin, W. W. (2004), 'Factors motivating software piracy: a longitudinal study', *IEEE Transactions on Engineering Management* **51**(4), 414–425.

Maltz, E. & Chiappetta, V. (2002), 'Maximizing value in the digital world', *Sloan Management Review* **43**(3), 77–84.

Moulin, H. (1994), 'Serial cost sharing of excludable public goods', *Review of Economic Studies* **61**, 305–325.

MPAA (2008), Motion Picture Association of America: Anti-piracy. `http://www.mpaa.org/piracy_OptDisk.asp`. Accessed 02/22/2008.

Nascimento, F. & VanHonacker, W. R. (1988), 'Optimal strategic pricing of reproducable consumer products', *Management Science* **34**(8), 921–937.

Newman, M. E. J. (2003), Random graphs as models of networks, *in* S. Bornholdt & H. G. Schuster, eds, 'Handbook of Graphs and Networks', New York: John Wiley & Sons, pp. 35–68.

Norman, P. (2004), 'Efficient mechanisms for public goods with use exclusions', *Review of Economic Studies* **71**(4), 1163–1188.

Ordover, J. A. & Willig, R. D. (1978), 'On the optimal provision of journals qua sometimes shared goods', *American Economic Review* **68**(3), 324–338.

RIAA (2008), Recording Industry Association of America: Piracy, online and on the street. `http://www.riaa.com/physicalpiracy.php`. Accessed 02/22/2008.

Serizawa, S. (1999), 'Strategy-proof and symmetric social choice functions for public good economies', *Econometrica* **67**(1), 121–145.

Sundararajan, A. (2004), 'Managing digital piracy: Pricing and protection', *Information Systems Research* **15**(3), 287–308.

Varian, H. R. (2005), 'Copying and copyright', *Journal of Economic Perspectives* **19**(2), 121–138.