

# Social Sharing of Information Goods: Implications for Pricing and Profits

Michael R. Galbreth

Moore School of Business, University of South Carolina

Bikram Ghosh

Moore School of Business, University of South Carolina

Mikhael Shor

Department of Economics, University of Connecticut

Forthcoming in *Marketing Science*, 2012

## Abstract

Social sharing of information goods—wherein a single good is purchased and shared through a network of acquaintances such as friends or coworkers—is a significant concern for the providers of these goods. The effect of social sharing on firm pricing and profits depends critically on two elements: the structure of the underlying consumer network and the mechanism used by groups to decide whether to purchase at a given price. We examine the effect of social sharing under different network structures (decentralized, centralized, and complete), which reflect a range of market conditions. Moreover, we draw from the mechanism design literature to examine several approaches to group decision making. Our results suggest that a firm can benefit from increased social sharing if the level of sharing is already high, enabling a pricing strategy targeted primarily at sharing groups rather than individuals. However, the point at which sharing becomes marginally beneficial for a firm depends on both the distribution of group sizes (which derives from the network structure) and the group decision mechanism. Additional insights are obtained when we extend the model to capture homophily in group formation and the potential that a subset of consumers will never share for ethical reasons.

**Keywords:** information goods; social networks; pricing; graph theory; cost sharing

# 1 Introduction

Sharing of information goods is not a new phenomenon, but until recently its scope has been limited. Historically, information goods such as music, stock quotes, and university course content were inexorably attached to the CDs, newspapers, and classrooms through which they were delivered. Today, many information goods can be digitized and separated from their physical media, and the non-rivalrous nature of the digital format allows for broader sharing among multiple simultaneous users. As noted by Gopal & Sanders (1998), Varian (2000), Peitz & Waelbroeck (2006), and others, it is now easy for groups of consumers—whether they be family members, friends, or colleagues—to share information goods, with each potentially contributing to its cost.

For example, while single newspaper subscriptions have long been shared within families, the sharing of a password for a subscription news Web site is much less constrained (Hwang et al. 2009). Sharing of passwords to online accounts has been documented in music services (MusicAlly 2005), the real estate Multiple Listing Service (Messmer 2009), financial analysis (Pritchard 1998, Burton-Taylor 2009), online supply chain exchanges (Wilson 2008), and sports services (Philadelphia Inquirer 1997). In the case of online music services, Stubblefield & Wallach (2000) note that “of likely concern to the recording companies, users can trivially share passwords and... could easily pool their money, purchase individual copies of music CDs... and then the whole cartel can access the music.” Similarly, many financial analysts fear that putting stock tips online would lead to lost revenue, as “too many small investors would club together to share a subscription to the site” (Pritchard 1998, p. 2). Indeed, the *Financial Times* recently sued a major equity firm for alleged sharing of login information among colleagues (Herman 2009), and specialized software has been developed to help combat such activities (Messmer 2009).

In this paper we examine the impact of social sharing of information goods, such as subscription Web sites, on the pricing and profits of the producer of the good. We define social sharing as sharing of goods *through a network of social acquaintances* (coworkers, friends, etc.). This is in contrast to anonymous sharing, exemplified by several well-publicized peer-to-peer networks in which goods are relayed *through an anonymous online*

*proxy*. Anonymous sharing over peer-to-peer networks has increasingly been reduced via technical means (see, for example, Smith & Fowler 2011), while social sharing is more difficult to combat. The result is that social sharing, while perhaps receiving less media attention, is substantially more damaging to firms—for example, the value of financial information illegitimately accessed through shared accounts each year is estimated at eight billion dollars (Burton-Taylor 2009), more than twice the annual losses due to movie piracy (MPAA 2008). In addition, social sharing differs from anonymous sharing in terms of how the costs of the acquired goods are divided among those who share it. While free riding is a potential issue in anonymous sharing networks,<sup>1</sup> social sharing groups can easily solve free riding issues by excluding non-contributors.

Social sharing of information goods need not be troublesome from a firm’s perspective. This is particularly true when firms are able to price discriminate between sharing groups and individual consumers. Owners of information property have often confronted sharing with substantially higher prices. Journals are typically sold to libraries at much higher prices than individual subscriptions (Ordovery & Willig 1978). Similarly, intertemporal price discrimination was used by firms to charge video rental stores many times the price for movies as they charged home consumers (Mortimer 2007). These higher prices partially appropriate multiple users’ values of the shared product (Liebowitz 1985). However, while libraries and video rental stores are easily distinguished from individuals, in contexts such as password sharing among colleagues, groups of consumers can essentially form their own private libraries, rendering such price discrimination impossible.

We take a graph theoretic approach to modeling social sharing group formation, using well-established social network structures to represent potential connections among consumers. A consistent finding across theoretical, behavioral, and empirical studies is that the ‘ring’ and the ‘star’ structures—representing decentralized and centralized networks, respectively—are appropriate stylized representations of the types of social networks used for sharing information goods. We consider both of these structures, as well as a ‘complete’ network, in which each person is equally likely to form a direct link with every

---

<sup>1</sup>It should be noted, however, that technical strategies such as rewarding users who share with faster download speeds, sometimes to the complete exclusion of non-sharers (Hughes et al. 2005), have reduced free riding in peer-to-peer settings to as little as five percent in some cases (Andrade et al. 2005).

other person. Next, using the mechanism design literature, we consider two possible decision mechanisms within groups—where group members truthfully reveal their valuations in an attempt to raise enough funds (efficient mechanism) and where they actively seek to understate their valuations, insisting that all group members contribute equally (equal-pay budget balancing mechanism). The decision-making process employed might, for example, be driven by the nature of the social relationship between group members (e.g. family members vs. coworkers).

Our main results and their implications for marketing managers can be summarized as follows. First, although managers might typically assume that increased social sharing is detrimental to profits, we show that this need not be the case if they set prices accordingly. In fact, regardless of the underlying network structure, if sharing is already widespread, then additional sharing always increases firm profits. This is because higher levels of sharing result in a consolidation of the market, making it easier for the firm to target its price towards sharing groups rather than individuals. Once prices can be targeted to groups, additional sharing (i.e. more group formation) increases firm profits, and thus efforts to combat it might not be justified.

Second, we show that understanding the underlying structure of the consumer network can help managers determine the likely impact of sharing and their optimal response to it. If the consumer network is centralized, then the market will consolidate more quickly as sharing increases, and the switch to group pricing occurs at lower levels of sharing activity. Thus, firms facing a centralized network of consumers are more likely to benefit from sharing than those facing a decentralized network. This differential effect of network structure is particularly useful since empirical studies have identified market conditions under which a particular network structure is likely to evolve. For example, the findings of Baker & Faulkner (1993) suggest that a centralized (decentralized) network may evolve in a market wherein there is a low (high) need for concealment of sharing activities. Since the need for concealment may be a function of the strength of a particular market's anti-piracy laws, our results suggest that managers should consider the stringency of the regulatory environment when determining their response to social sharing of their

products.

Third, we show that managers should seek to understand how their consumers collaborate within groups to decide whether to purchase a good and share it. We find that firms are more likely to benefit from sharing when groups use an efficient decision mechanism, such as when family members or friends willingly reveal their true valuations to each other and cooperate to make the purchase. On the contrary, groups of consumers less inclined to be truthful about their valuations might use an equal-pay mechanism, for example financial professionals sharing access to expensive reports or real estate agents sharing multiple listing service access. Our results imply that efforts to combat social sharing are more suited to the latter examples. Finally, the interaction of the network structure and the group decision mechanism add several nuances to our main findings.

Our base model assumes, following the graph theoretic literature, that all possible connections between consumers are equiprobable, and that sharing always occurs across realized links. Additional insights are obtained in §6 from several relaxations of these assumptions. First, we consider homophily, where a group is more likely to contain similar consumers than dissimilar ones. While correlation among group members' valuations reduces the benefit of statistical averaging of those valuations, we show that, surprisingly, sharing in homophilic groups may nevertheless enhance profits. Second, we examine the effect of a subset of consumers who refuse to share despite opportunities to do so, perhaps due to ethical concerns. We show that the growth of this ethical consumer segment may actually decrease profits, as it constrains the firm from pricing optimally for groups who do share. Across the core model and extensions, from a methodological perspective we show that firm profits depend on the *distribution of group sizes* arising from the specific network structure, as well as the group decision mechanism. Thus, optimal pricing can be pursued even without in-depth knowledge of the network structure by instead focusing on the proportion of consumers who share the good and the number of consumers with whom they share.

To our knowledge, this is the first paper to analyze the pricing and profit implications of sharing information goods within groups while explicitly considering the social network

structures that give rise to the groups and the group decision mechanisms that determine whether a purchase is made. In formulating our model and its extensions, we draw on a broad set of literatures, as described in the next section. Our unique approach to this problem enables us to provide new insights for managers who are interested in how social network structures and group dynamics should inform their firm's reactions to sharing among consumers. We also provide a novel modeling framework for marketing researchers by using graph theory to incorporate social networking concepts into pricing models.

## 2 Related Literature

This paper relates to the extant marketing and information systems literatures on sharing digital products. Much of this research has focused on the network effects of sharing. Conner & Rumelt (1991) and Shy & Thisse (1999) show that copying of software may have a positive effect on a monopolist's profit when markets exhibit network externalities, such as word-of-mouth information transfer from a larger installed base (see Givon et al. 1995). Sharing may also allow for product sampling before purchase (Bhattacharjee et al. 2003). More recently, Jain (2008) analyzes digital piracy in a competitive market. With competition, sharing by price-sensitive consumers results in higher equilibrium prices for all firms, which may outweigh any negative effects of piracy. While the above models require either network effects on the consumer side or competition-softening effects on the firm side to obtain positive effects of sharing for the firm, in this paper we show that profit gains can be possible in a monopolist setting as well, even without network effects.

Also related to this paper are studies on deterrents of digital piracy, such as improved functionality of legal Web sites, perceived risk from piracy, and concern for social norms (Sinha & Mandel 2008), or drivers of piracy, such as higher prices (Cheng et al. 1997). In one of the first papers to address group formation to purchase information goods, Gopal & Sanders (1998) assume that consumers deterministically form into identical groups of optimal size given the pricing and control decisions of the firm. In reality, heterogeneous groups are likely to form, depending probabilistically on the underlying social network structure, as described in our model setup below.

Our work also relates to the economics literature, where prior research has shown that sharing of information goods can lead to higher profits for firms (Bakos et al. 1999, Varian 2005). Peitz & Waelbroeck (2006) provide a comprehensive survey of the economics literature on illicit sharing. Most closely related to our work is Bakos et al. (1999), who show that firms can profit from consumer sharing because sharing “reshapes demand” by summing the valuations of consumers in groups. Since the sum of valuations exhibits less variance than the valuations themselves, this enables firms to increase profits. We show that this profit effect depends on both the structure of the underlying social network and the decision mechanism employed within groups. The reduction of variance within groups is accompanied by an increase in variance across groups as groups of different sizes form. The implications of sharing depend on which of the two effects dominates. Building on Bakos et al. (1999), our paper is the first to consider explicitly how sharing groups might arise. Moreover, whereas Bakos et al. (1999) make the common assumption that groups fully aggregate members’ valuations, we also consider inefficient decision mechanisms, which may more accurately reflect many real-world groups. Finally, while Bakos et al. (1999) assume that all links between consumers are equally likely, we also consider the case of homophily, or preferential attachment between people with similar valuations, as well as the possibility that some subset of consumers will not consider sharing at all.

In the following section we develop our model, describing in detail our approaches to modeling both consumer network structures and group decision mechanisms.

### 3 Model

We consider a monopolist producer of an information good faced with consumers who may share the good via a social network. We begin with a simple example to highlight some of the intuition of the results that follow. Consider a marketplace consisting of three consumers who value a particular good at \$11, \$12, and \$16. In the absence of sharing, the profit maximizing firm would set a price of \$11, earning a profit of \$33 (ignoring production costs). If sharing is a possibility, there are five different network structures that could be realized. First, each consumer can be independent, as above.

Second, the three consumers could form a single group, purchasing and sharing a single copy among them. In this case, the firm could charge a maximum of \$39, the sum of all consumer valuations, in the best case scenario where a group fully aggregates the values of its individual members. Thus, selling to a single group of consumers allows for higher profits than selling to each consumer individually.

Lastly, three different configurations allow two of the three consumers to form a group: (i) (\$11,\$12),\$16 ; (ii) (\$12,\$16),\$11 and (iii) (\$11,\$16),\$12. If (i) is realized, the firm selects a price of \$16 at a profit of \$32, as the group purchases a single good to share. The corresponding price and profit for (ii) and (iii) are (\$28,\$28) and (\$27,\$27) respectively. In this example, any of these three consumer configurations result in lower profits than selling to independent consumers. Here, while each group aggregates its members' values, the heterogeneity in group sizes reduces profits.

The above example is oversimplified in three important ways. First, in reality the firm cannot observe individual consumer valuations. Secondly, the firm cannot observe the sharing groups that have formed—instead, it must set a price that is optimal given the probabilistic formation of all possible groups. Third, group valuations need not always fully reflect the sum of members' valuations. We address the first point by modeling each consumer's valuation of the good as a random variable. Our approach to the the second and third points is described in the following subsections, where we discuss and formalize both the social networks that give rise to sharing groups and the decision mechanisms that groups use to make purchase decisions.

### 3.1 Consumer Networks

A social network is a collection of people and a set of possible social ties between them, often represented by a random graph (Newman et al. 2002). We consider the ring, star, and complete network structures. These represent, respectively, social networks that are highly decentralized, highly centralized, and complete (in that each person is equally likely to form a direct link with every other person).<sup>2</sup> While any of these structures may

---

<sup>2</sup>Formal definitions of these structures are provided below. Our focus here is simply on grounding our choices of network structures within the broader literature.

theoretically evolve, following the larger game theoretic and graph theoretic literature, our primary focus in the subsequent analysis is on the ring and star network structures, with the complete network examined separately as an extension (§6.1). The rationale for this focus is as follows. In a seminal paper, Bala & Goyal (2000) show that only two network structures—the ring and the star—are strict Nash equilibria of strategic network formation. Other models of strategic formation also converge to either a star (Feri 2007) or a ring (Deroian 2009). Ring and star structures receive not only theoretical support, but are consistent with both behavioral studies and many real world networks. Falk & Kosfield (2003) experimentally confirm much of the theoretical results in Bala & Goyal (2000), highlighting conditions under which a ring-like structure is likely to evolve. Goeree et al. (2009) experimentally show that some individuals in a network have a lower opportunity cost of networking and hence are more likely to be central nodes. As such they demonstrate that heterogeneity of agents and information asymmetry are primary facilitators in the evolution of star networks. In a large-scale empirical study, Yahoo! Research find that a significant fraction of online social networks are star topologies (Kumar et al. 2010). In another empirical study, Baker & Faulkner (1993) find that groups formed to conduct illicit activities favor decentralized (e.g. ring) network structures to allow for concealment and increase the difficulty of uncovering a group leader. When there is less need for concealment, Baker & Faulkner (1993) find that a centralized, star network might be employed. Overall, the behavioral and empirical literatures strongly support both the ring and star topologies as likely social network structures from which sharing groups might arise.

An important implication of the work described above is that not *all* direct links between *all* consumers have a positive probability, as is assumed by the complete network topology. Rather, many consumer links are indirect, made via intermediate consumers. Social network studies largely confirm this finding, suggesting that direct links might be quite unlikely between certain individuals due to their associations with different communities (Girvan & Newman 2002, Newman 2001). However, we remark that a positive probability of all direct links between all individuals is theoretically possible.

Such a ‘complete’ graph derives from the early work of Erdős & Rényi (1960), and it has some support from the economics literature (e.g. Jackson & Wolinsky 1996). Thus, we examine the complete graph structure as a codicil to our main results in §6.1.

### Formal Definitions of Consumer Network Structures

A random graph  $\Gamma = \{\mathbf{N}, \mathbf{E}, q\}$ , prescribes a set of consumers,  $\mathbf{N} = \{1, \dots, N\}$ , and a set of edges,  $\mathbf{E}$ , representing possible linkages between consumers, each occurring independently with probability  $q$ .<sup>3</sup> Define  $h$  as a set of the sizes of connected components in a realization of  $\Gamma$ , representing the partition of consumers into groups. Therefore,  $h$  is an integer partition of  $N$ . For example,  $h = \{3, 3, 2\}$  represents eight consumers partitioned into three groups—two with three members, and one with two members. The two structures of random graphs are defined by the sets of vertices:

- Decentralized (Ring):  $\mathbf{E} = \{(1, 2), (2, 3), \dots, (N - 1, N), (N, 1)\}$
- Centralized (Star):  $\mathbf{E} = \{(1, N), (2, N), \dots, (N - 1, N)\}$

Figure 1(a) provides an example of the decentralized ring structure for an economy of eight consumers. In the sample realization, represented by darkened edges, the firm faces four groups: two consisting of two consumers, one consisting of three consumers, and one singleton. Figure 1(b) presents an example of the centralized star structure, which is comprised of a single entity with whom all other nodes have a potential connection. Since peripheral nodes may only share through the central node, at most one multi-person group may arise under this structure (a five-consumer group is shown in Figure 1(b)).

### Expected Group Size Distribution

Given  $q$  and the number of consumers in the market  $N$ , we define  $\tau_k$  as the expected number of groups of size  $k = 1, 2, \dots, N$ . Note that  $\sum_{k=1}^N k\tau_k = N$ . Furthermore, when

<sup>3</sup>The assumption of independent  $q$  is relaxed in §6.2, where homophily in link formation is considered. Furthermore, we note that in reality  $q$  may be affected by the firm’s pricing decision (Cheng et al. 1997). Linking  $q$  to price would require an ad hoc assumption regarding the functional form of the linkage. Similar to related work in this area (e.g., Bakos et al. 1999), for parsimony such a linkage is not modeled here. A qualitative discussion of the impact of a link between  $q$  and price in our context is provided in §7.

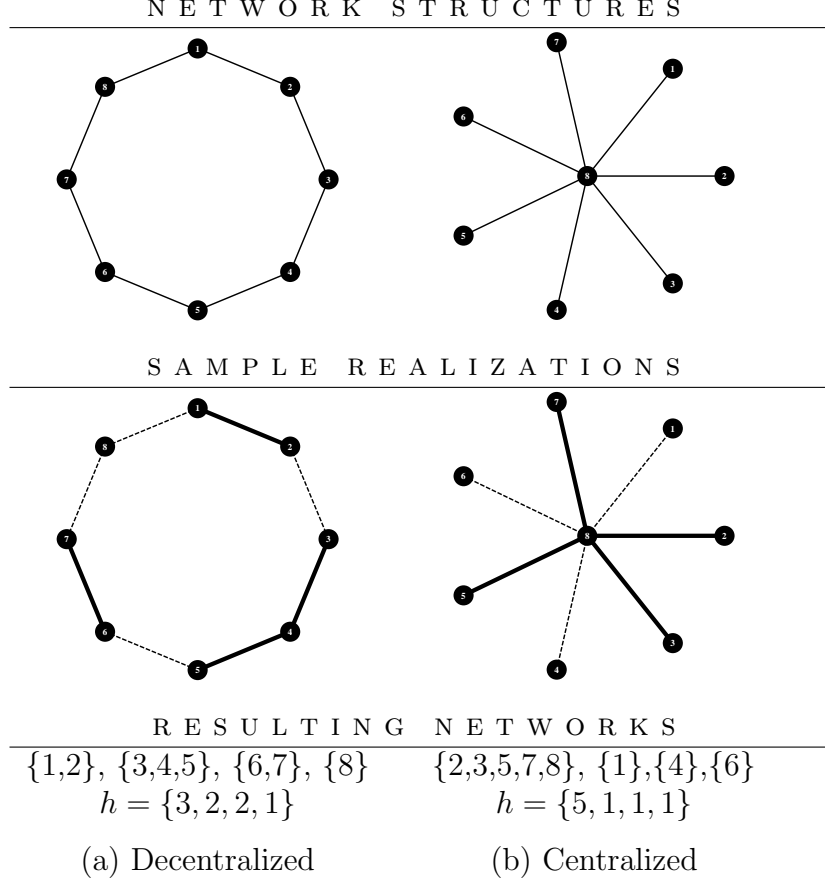


Figure 1: Network structures, with example realizations

$q = 0$ ,  $\tau_1 = N$  (no groups form) and when  $q = 1$ ,  $\tau_N = 1$  (all consumers form into a single group). The following proposition derives  $\tau_k$  for the decentralized and centralized network structures.

**Proposition 1.**

*For the decentralized network:*

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

$$\tau_N = Nq^{N-1}(1-q) + q^N$$

*For the centralized network:*

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

$$\tau_1 = (N-1)(1-q) + (1-q)^{N-1}$$

The proofs of all results can be found in the Supplementary Appendix. We first identify how group sizes change with  $q$ , which we summarize in the following corollary:

**Corollary 1.1.**

$$\begin{aligned}
\frac{d\tau_1}{dq} < 0, \frac{d\tau_N}{dq} > 0 & \quad \text{for all networks} \\
\frac{d\tau_k}{dq} > 0 \Leftrightarrow q < \frac{k-1}{k+1} & \quad k \in \{2, \dots, N-1\} \quad \text{for the decentralized network} \\
\frac{d\tau_k}{dq} > 0 \Leftrightarrow q < \frac{k-1}{N-1} & \quad k \in \{2, \dots, N-1\} \quad \text{for the centralized network}
\end{aligned}$$

For both networks, the expected number of multi-person groups,  $\tau_k$  ( $k \geq 2$ ), increases in  $q$  up to a threshold and then decreases. This implies that at low levels of sharing activity, the market remains fairly fragmented, consisting of individual consumers and small groups. Ceteris paribus, a firm's pricing problem is more difficult when faced with such a heterogeneous market. For  $q$  above the threshold,  $\tau_k$  ( $k < N$ ) decreases with  $q$ , reflecting consolidation into larger groups, with fewer individual consumers. Corollary 1.1 establishes that, for  $k < N - 2$ , this threshold value of  $q$  is higher in the decentralized network than in the centralized one. Thus, market consolidation into larger groups occurs at lower levels of sharing likelihood in the centralized network. This plays a critical role in the firm's ability to raise prices in response to sharing activity, as shown in the analysis in §4.

**3.2 Group Decision Mechanisms**

A group of consumers wishing to share an information good must make two related decisions: whether to purchase the good and how to divide its cost. Once the good is purchased, it becomes a 'public good' in the sense that all group members benefit equally, regardless of the amount paid. The behavioral and theoretical economics literatures both examine how individuals contribute to acquire a public good.

A potential issue with eliciting contributions toward the acquisition of a public good is the incentive for individuals to understate their true valuations of the good. However, this issue does not necessarily exist in all sharing groups. For example, close friends may willingly reveal their valuations and simply purchase the good whenever their values collectively exceed its price. In other words, they will employ an efficient mechanism

(Clarke 1971). On the contrary, members of groups without such close ties may understate their true valuations, rendering the decision mechanism inefficient. When efficiency cannot be attained, the group mechanism can instead ensure ‘budget balancing,’ i.e. the collection of payments which exactly equal the price. We consider both the efficient and an equal-pay budget balancing mechanism, as described in the following subsections.

### **3.2.1 Efficient Mechanism**

Consider first the case where group members either willingly reveal their valuations or, alternately, where a group devises a mechanism (e.g., Groves-Ledyard) which leads to truthful revelation in equilibrium. In both cases, the group’s decisions are efficient, in that a purchase occurs precisely when the sum of all individual valuations exceeds the price. Such an assumption is appropriate, for example, when group members are willing to cooperate with others within the group to acquire the good. While honest revelation of values and cooperation in the absence of a mechanism may seem unlikely in group decision making, experiments in economics and psychology suggest that it is more likely than commonly assumed.

Since Dawes et al. (1977), numerous experiments have examined people’s contributions towards the purchase of a public good. For example, groups may exhibit cooperative behavior and significant contribution to public goods can occur without any mechanistic infrastructure or extrinsic incentives, motivated by a sense of altruism and reciprocity, even among strangers (Croson 2007). Social ties among group members further raise contributions, even among people predisposed to acting selfishly (VanVugt & DeCremer 1999), as do experience and repeated interaction (Ostrom 2000) and face-to-face communication (Ledyard 1995, Ostrom 2000). Many of these moderating psychological processes may be present when in groups of family and friends. Furthermore, social and economic sanctioning of non-cooperative members also increase contributions (Gachter 2000). Overall, experimental studies have found quite high find rates of successful purchase when it is efficient to do so, sometimes approaching 90% (e.g., Bagnoli & McKee 1991).

### 3.2.2 Budget Balancing Mechanism

Although efficiency is a reasonable reflection of reality in a surprising range of real world contexts (and a common assumption in economic models), it is clearly not achievable in all groups, particularly when cooperation levels are low. We also consider the other side of the theoretical coin, a mechanism which ensures *budget balancing* but may lead to an understatement of group members' valuations, and thus an *inefficient* choice not to purchase despite the sum of groups members' values 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). The mechanism is also intuitively simple and quite easy to implement, involving each willing group member paying an equal share of the price. We refer to such a mechanism as 'equal-pay budget balancing.' For example, two friends will acquire a good at a price of \$20 if each of them values it at more than \$10, or if one of them values it at more than \$20. More formally, a group of size  $n$  purchases a good at price  $p$  if there exists an  $m \leq n$  such that  $m$  people value the good at least  $p/m$ . In practice, the group asks if each member is willing to pay  $p/n$ . If all agree, the good is purchased and each pays an equal share. Otherwise, the group checks if  $n - 1$  people are willing to 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 \$8 and \$16 would not purchase a good priced at \$20 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 summary, our efficient mechanism is consistent with experimental evidence on sharing in social networks. On the other hand, many less close-knit groups adopt a norm of equal payment by each contributing member, despite heterogeneity in values (Levati et al. 2007). This is precisely captured by the equal-pay budget balancing mechanism described above. Thus, when selecting stylized representations of group decision making, the efficient and the equal-pay budget balancing group mechanisms appear appropriate.

### 3.3 Costless Sharing

Following the larger literature on sharing of information goods (Varian 2005, Bakos et al. 1999), we assume that sharing is without degradation of quality (i.e. is costless). This is reasonable given the fact, noted by Gopal et al. (2006) and others, that digital products are fundamentally non-rivalrous in the sense that sharing them does not reduce their utility. The example of password sharing seems to be well-suited to this assumption, and researchers have suggested numerous other cases where costless sharing is reasonable, for example social sharing of software (Varian 2000).<sup>4</sup> We note that a separate but related research stream addresses costly sharing, often based on Digital Rights Management (DRM) initiatives, which attempt to impose a cost on duplication (Desai et al. 2011).

In the next section we analyze the firm's pricing problem for both centralized and decentralized network structures, first under the efficient group decision mechanism, and then under the equal-pay budget balancing mechanism.

## 4 Analysis

### 4.1 Efficient Mechanism Analysis

Each consumer's valuation of the good is independently distributed with distribution function  $F(v)$ . We begin with the efficient mechanism, so that a group of size  $k$  purchases the good if the sum of its members' valuations exceeds the price. Therefore, the distribution of the group's valuation is given by  $F_k(v) = F(v_1) * \dots * F(v_k)$ , i.e. the  $k$ -fold convolution of  $F$ .

In the absence of sharing, the monopolist maximizes  $\Pi(p) = p(1 - F(p))$ , the price times the chance of purchase. In the presence of sharing, the firm's problem is more complex. A network structure yields a probability distribution over possible sizes of groups. The resulting probability distribution is thus a random convex mixture of convolutions. However, the following claim shows that the network structure enters the profit function only through the expected number of groups of each size. Thus, the profit function for

---

<sup>4</sup>As noted by Peitz & Waelbroeck (2006), when sharing is among social acquaintances, supporting documentation or other supplemental information on the product can be obtained, minimizing the utility differences between copies and originals.

the monopolist can be significantly simplified.

**Claim 1.** Define  $F_k(v)$  as the distribution function of the valuation of a group of size  $k$ , and  $\tau_k$  as the expected number of groups of size  $k$ . The firm's expected profit is given by:

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

In words, 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. This establishes that firm profits depend on the social network structure only to the extent that different structures give rise to different values of  $\tau_k$ . This allows us to represent sharing 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.

To derive specific results on firm profits, we first must specify a distribution for consumer valuations. We adopt the exponential distribution for  $v$ . The exponential distribution is desirable since the sum of exponentials is given by the Gamma distribution, which can accommodate both fat tails and asymmetry in distribution. The “fat-tail” property of reservation values is supported by the marketing literature. For instance, Sonnier et al. (2007) document that willingness to pay has substantial mass in the tails of the support. Finally, though it is common in stylized models to assume uniformly distributed consumer preferences, the uniform distribution is not closed to convolutions, making our analysis considerably less tractable. Additionally, as the uniform distribution is bounded from above, a firm can easily monitor sharing since only groups would be willing to pay above the upper bound. Results similar to those reported were obtained for several parameterizations of the gamma, chi-square, and Weibull distributions (available from the authors upon request).

For the exponential distribution, the density and distribution of the  $k$ -fold convolution are given by:

$$f_k(v) = \frac{v^{k-1}}{(k-1)!} e^{-v} \text{ and } F_k(v) = 1 - \sum_{i=1}^k \frac{v^{i-1}}{(i-1)!} e^{-v} = 1 - \sum_{i=1}^k f_i(v) \quad (2)$$

Due to the non-transcendental nature of this and similar functions involving convolutions of random variables bounded from below, this expression for  $F_k(v)$  precludes the derivation of an explicit closed-form solution for the optimal price.<sup>5</sup> However, the results are quite stable for large  $N$ , enabling us to optimize numerically, using  $N = 150$  as a representative market size (the results are very similar, and all insights discussed below persist, when  $N > 150$ , although the computation of optimal profits and prices becomes increasingly time-intensive). Importantly, we note that this numerical approach does not sacrifice generalizability in our context since the only parameter in the model other than  $N$  is  $q$ . Thus, the optimal profit and price dynamics for large  $N$  are completely characterized in general by the figures in this section. Moreover, the initial profit impact of sharing at low  $q$ , as well as its eventual impact for high  $q$ , are established analytically in Claim 2.

**Claim 2.** (a) For  $q$  sufficiently low, profit is increasing in  $q$  in the decentralized network and is decreasing in  $q$  in the centralized network. (b) For  $q$  sufficiently high, profit is increasing in  $q$  in the centralized network.<sup>6</sup>

Results for the decentralized network are illustrated in Figure 2. In panels (a) and (b), we see that profits and prices are little changed in the decentralized case for low values of  $q$  (though profits are increasing—see Claim 2). This can be attributed to the fact that, following Proposition 1, the number and size of multi-person sharing groups gradually increases with sharing activity. A firm considering higher prices faces a tradeoff between capturing added surplus from these multi-person groups and losing sales from individuals (that is, groups of size one, the number of which strictly decreases in  $q$ , as shown in panel (b)). For nearly all levels of sharing activity, the firm is able to increase prices slightly to account for group formation and avoid a negative profit impact. In other words, prices remain relatively flat when sharing activity is low—a finding that is consistent across all of our analysis. However, at very high levels of sharing (high  $q$ ), most consumers are

---

<sup>5</sup>Explicit results can be obtained if we consider a distribution on an unbounded domain. However, allowing for negative valuations of a product is not a reasonable reflection of reality, so the bounded distribution is used.

<sup>6</sup>The decentralized profit function presents mathematical difficulties that preclude a complete analysis of its behavior for high  $q$ . However, in the Supplemental Appendix we verify that profits are increasing for high  $q$  in the decentralized case as well.

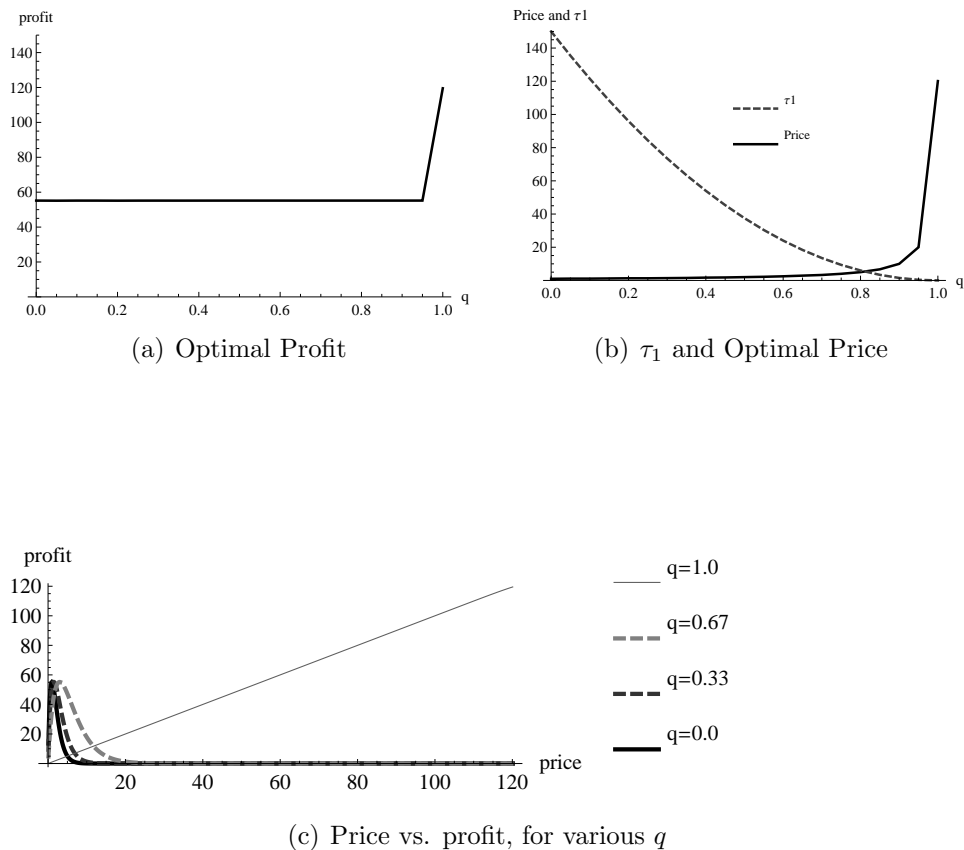


Figure 2: Profits and prices in decentralized networks as a function of sharing probability.

members of larger groups, allowing the firm to raise prices aggressively. This represents a transition from pricing geared more toward individuals to pricing for groups. Overall, we find that when sharing networks are highly decentralized, the presence of sharing is likely to have very little effect on prices or profits, though both do increase slightly with increased sharing. Profit functions presented in panel (c) further demonstrate the sluggish change in the optimal price for most levels of  $q$ .

Results for the centralized network are illustrated in Figure 3. Unlike in the decentralized network, low levels of sharing activity decrease firm profits. Whereas the decentralized network allows many small groups to form, the centralized network allows for a maximum of one sharing group as a central node offers the only opportunity for consumers to connect. At low levels of sharing activity, the firm mostly disregards this sharing group and chooses an optimal price for individual consumers. This lack of a pricing response to sharing activity is detrimental to firm profits, since one of the ‘individual consumers’ is actually a group in which consumers share a single purchase. As sharing

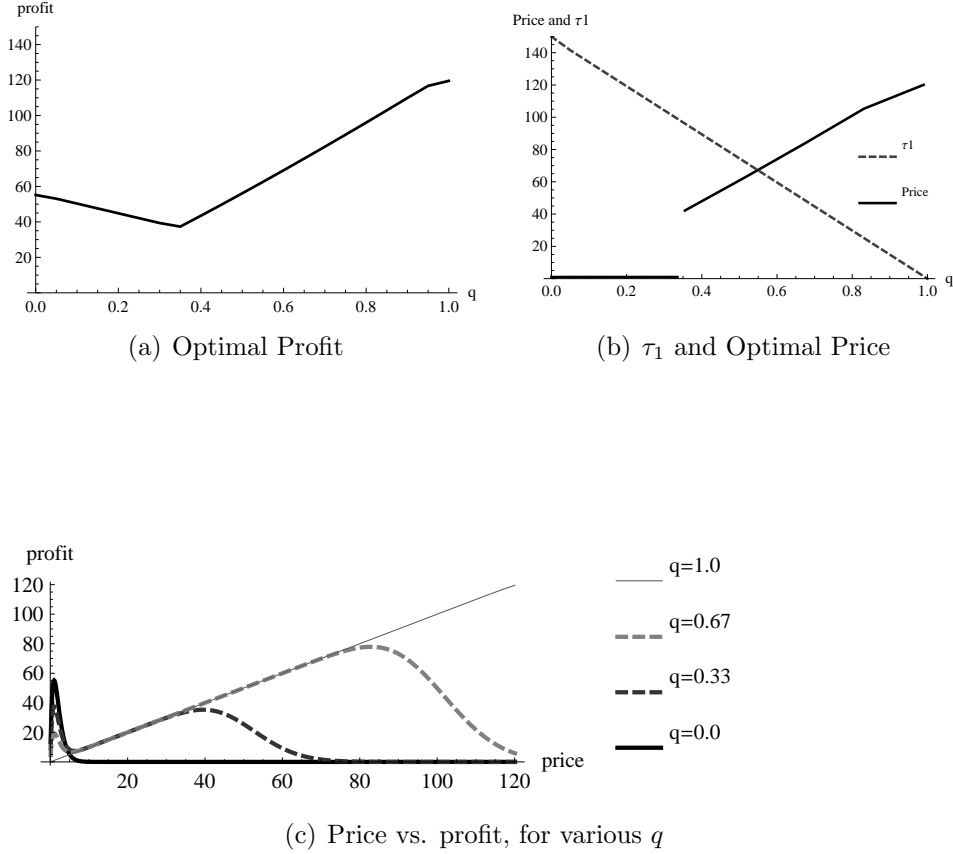


Figure 3: Profits and prices in centralized networks as a function of sharing probability.

increases, the size of the centralized group eventually makes it optimal for the firm to tailor its price to this group at the possible exclusion of individual consumers.

There is a threshold level of sharing in the centralized network above which the firm switches to group pricing. As established analytically in Claim 2, firm profits are increasing in  $q$  for sufficiently high  $q$  under this network structure. Profits in the centralized network at various levels of  $q$  are illustrated in Figure 3(c). Unlike under the decentralized structure, profits as a function of price in this case are double-peaked for  $0 < q < 1$ , representing the profits from pricing for individual consumers, optimized at the first peak, and pricing for the group, optimized at the second peak.<sup>7</sup> Lemma 2 in the Supplementary Appendix establishes that there do not exist any other local maxima. As  $q$  rises, the first peak becomes lower, as fewer unconnected consumers remain, and the second

<sup>7</sup>The single-peaked profit function in the decentralized case (Figure 2(c)) is due to the “non-discreteness” in optimal pricing for a decentralized network. Specifically, with a decentralized network the firm is not simply choosing between pricing for unconnected individuals and pricing for a single group (two discrete choices), but instead faces many different sizes and combinations of groups. For this reason, prices adjust smoothly, targeting increasingly larger groups as  $q$  increases.

peak becomes higher, representing an increased chance of a larger centralized group. The point at which the firm switches to group pricing (the discontinuity in panel (b) of Figure 3) is precisely the value of  $q$  at which the second peak in panel (c) becomes higher than the first, and thus pricing for the group becomes more profitable.

The difference in firm pricing strategies between the two network structures is noteworthy. Recall that, as established analytically by Corollary 1.1, the market consolidates more quickly in a centralized network. The effect of this consolidation is clear when comparing Figure 2(b) and Figure 3(b). As shown in those figures, the faster consolidation enables a transition to group pricing at a much lower level of sharing activity when social networks are centralized as opposed to decentralized. This can also be seen from the change in the expected number of individual customers,  $\tau_1$ , in figures 2b and 3b above. Recall that this is a *general* result given that  $q$  is the only parameter in the model other than  $N$ , which does not affect the shape of the optimal price or profit curves.

To summarize the findings in this subsection, we highlight several new insights that have emerged from our analysis thus far. First, we find that low levels of sharing have little effect on optimal prices (Figures 2(b) and 3(b)). The profit effect of low levels of sharing depends critically on the network structure, with only centralized networks experiencing a profit decline from sharing (Figures 2(a) and 3(a), and Claim 2). In addition, the firm's shift to group pricing takes place at lower levels of sharing in a centralized network than in a decentralized one (Figures 2(b) and 3(b)). Finally, if  $q$  is high enough to justify group pricing, then firm profits are increasing in  $q$ —i.e., when sharing activity is already high, the firm prefers even more sharing activity at the margin (Figures 2(a) and 3(a), and Claim 2).

## 4.2 Budget Balancing Mechanism Analysis

As described in §3.2.2, we consider an equal-pay budget balancing mechanism, wherein a group purchases the good whenever there exist  $m$  members willing to pay at least  $p/m$ . Denote by  $v_{(i,k)}$  the  $i^{\text{th}}$  largest realized value in a group of size  $k$ . Then, the maximum a group of size  $k$  would pay is equal to

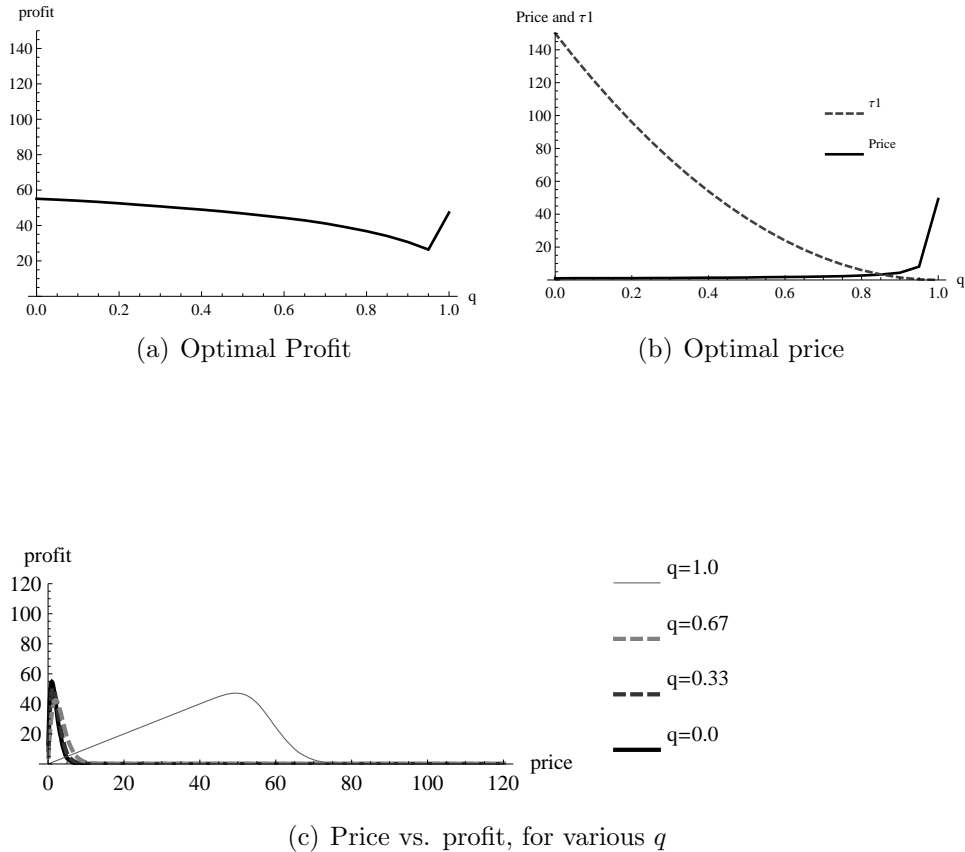


Figure 4: Profits and prices in decentralized networks with budget balancing decision mechanisms as a function of sharing probability.

$$F_k(v) = \max_{m \leq k} mv_{(m,k)}. \quad (3)$$

Clearly, this distribution of group valuations is stochastically dominated by the efficient decision mechanism. This is due to the inefficient aggregation of valuations in a budget balancing framework (also highlighted in the simple example in §3.2), which arises from the fact that surpluses are not transferred. That is, a consumer will not be able to acquire the good if her valuation is lower than her equal-portion allocation of the cost. This leads to relatively lower probability of purchase for the group relative to an efficient mechanism, where surpluses are transferred.

Despite this inefficiency, many of the insights from our efficient mechanism analysis carry over to the budget balancing case (see Figures 4 and 5). The marginal effect of  $q$  on profits, as with the efficient mechanism, is positive when  $q$  is high enough to justify group pricing. However, due to the lack of transfer of surpluses, this move to group

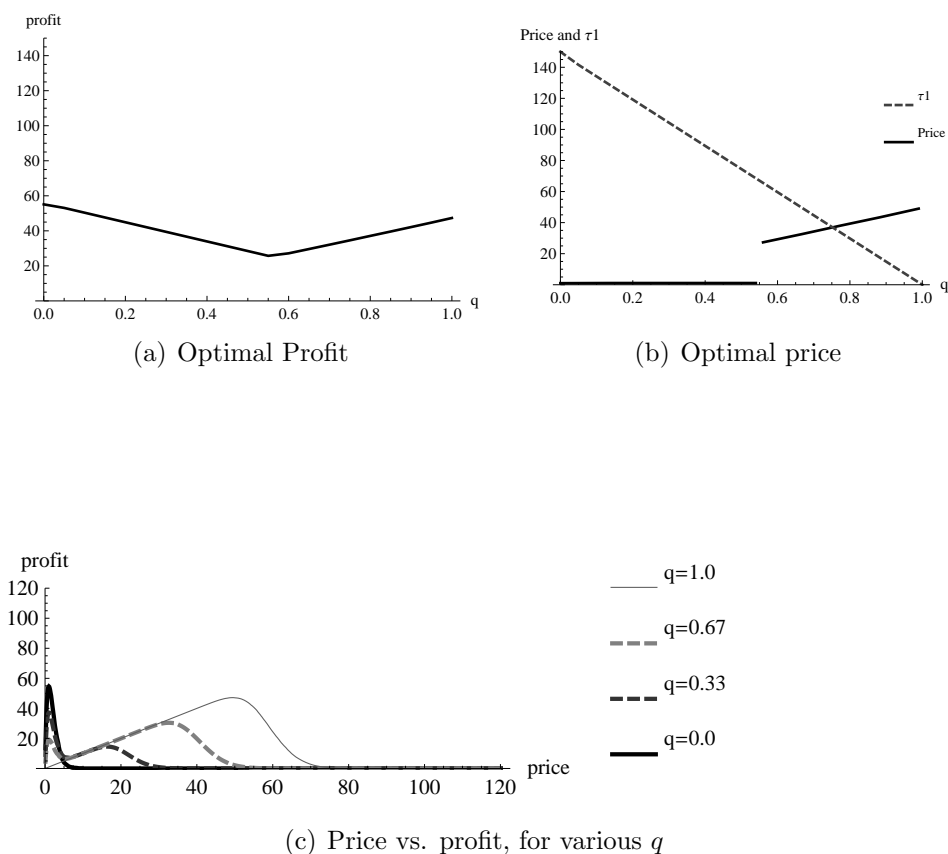


Figure 5: Profits and prices in centralized networks with budget balancing decision mechanisms as a function of sharing probability.

pricing occurs at a higher level of  $q$  with the budget balancing mechanism. Furthermore, all of the findings from the previous section that are solely driven by the network structure persist under this alternative decision mechanism—e.g., the threshold  $q$  at which the firm moves to group pricing is still lower in a centralized network than in a decentralized one, and the profit in a centralized network is still double-peaked.

One notable difference between the decision mechanisms is that, under the budget balancing mechanism, no sharing ( $q = 0$ ) is always preferred to any positive level of sharing. While this is an important distinction, we remark that firms rarely directly choose  $q$ , but rather, at best, may be able to influence the level of sharing within a specified range. Thus, the globally optimal value of  $q$  may be less germane than the impact of marginal changes in  $q$ . Taking this marginal perspective, we find that most insights persist—a threshold level of sharing still exists, above which profits are increasing in additional sharing activity, and below which profits are decreasing in sharing for the centralized network. However, we do find one important caveat for the decentralized

network—when consumer groups use an equal-pay budget balancing mechanism, profits are decreasing in sharing for most values of  $q$ . This is in sharp contrast to the efficient mechanism, under which profits are never decreasing in  $q$ . Thus, for managers of firms facing decentralized networks of consumers, the marginal profit impact of sharing depends critically on whether the group decision mechanism is efficient vs. equal-pay budget balancing.

Having established the main model results, before moving on to extensions we first provide an example of how our model can be applied to generate managerial guidance when real-world data on sharing activity is available.

## 5 Model Application: Calibration of $q$

We have established that the profit impact of increased sharing depends on the current level of  $q$ . In practice, firms might find the estimation of  $q$  to be challenging. However, our model can be used to calibrate  $q$  using data on the number of illegal copies in a particular industry. The expected proportion of illegal copies is given by  $1 - \sum \tau_k(1 - F_k(p)) / \sum k\tau_k(1 - F_k(p))$ .<sup>8</sup> One can calibrate  $q$  by setting  $p = p^*(q)$  in this expression (that is, assuming optimal pricing) and searching for the  $q$  that gives rise to the proportion of copies observed in practice. Consistent with our problem formulation, this approach assumes that  $q$  is not directly affected by price (see footnote 3). Thus, we remark that it yields a rough approximation of  $q$ , but nevertheless one that can provide general guidance in some real world contexts.

As an example, the percentage of software obtained through illegal copies in North America is estimated at 21% (BSA 2008). Assuming optimal pricing, the levels of  $q$  that gives rise to this proportion are  $q = 0.12$  (Decentralized, Efficient),  $q = 0.09$  (Centralized, Efficient),  $q = 0.13$  (Decentralized, Budget Balancing), and  $q = 0.08$  (Centralized, Budget Balancing). We can draw two implications from the numbers above. Locally, in all except the efficient decentralized model, profits are decreasing in  $q$  in the neighborhood of the

---

<sup>8</sup>This proportion is increasing in  $p$  for both the star and ring network structures. Note also that, under no sharing,  $\tau_1 = N$  and the above measure is zero; when  $\tau_N = 1$ , the measure reduces to  $(N - 1)/N$ . For the calibration, we use  $N = 150$ . As noted above, it can be confirmed that the results are stable for larger  $N$ .

values above (see Figures 2 through 5), indicating that, at present piracy rates, efforts to combat piracy are justified. At these levels of piracy, it is optimal for firms to price for individuals rather than groups, and a significant price increase is not an optimal response to illicit sharing in the North American software industry. This is consistent with anecdotal evidence that software firms are not dramatically increasing prices but are experiencing reduced profits due to sharing. However, a firm’s optimal response to illicit sharing might be very different in markets where piracy is much more prevalent, such as the estimated 80% piracy level in China and portions of Latin America (BSA 2008). The levels of  $q$  that give rise to such a high piracy rate are  $q = 0.67$  (Decentralized, Efficient),  $q = 0.35$  (Centralized, Efficient),  $q = 0.70$  (Decentralized, Budget-Balancing), and  $q = 0.56$  (Centralized, Budget-Balancing). In such a context, the effect of marginal changes in sharing activity on firm pricing and profits will clearly be quite different, since a shift to group pricing might be optimal at these levels of  $q$ , depending on the network structure.

These simple examples of software piracy are effective in demonstrating how our model might be applied to inform pricing strategies in a specific real-world context. A word of caution, however, is in order here. As with any model application, the quality of the outputs is limited by the quality of the inputs. In our calibration above, we implicitly assume that sharing group formation is confined within the physical boundaries of a particular country. While such an assumption is likely valid in our context of social sharing, it might not hold in some cases. In reality, firms might have more detailed data on the extent and geographic scope across which their products are shared. If so, they can use their data to perform their own model calibration following the example provided above.

Next, we examine several extensions to our main model, enabling us to provide additional insights into cases where a highly connected network structure exists, homophily is present in groups, or a subset of consumers who will never share exists.

## 6 Extensions

### 6.1 Complete Network

In this subsection, we consider a ‘complete’ network, where every consumer is potentially directly connected to every other consumer. Formally, given a random graph  $\Gamma = \{\mathbf{N}, \mathbf{E}, q\}$ , in a complete network, the set of edges is given by  $\mathbf{E} = \cup_{i < j} (i, j)$  for all  $i, j \in \mathbf{N} = \{1, 2, \dots, N\}$ .

Unlike the star and ring structures, nonasymptotic properties of complete networks are rarely obtainable in closed form. In our context, explicit solutions for  $\tau_k$  cannot be derived in closed form for the complete network,<sup>9</sup> but we can obtain them numerically. Asymptotically,  $\tau_k$  has a discontinuity at  $q = \frac{1}{N}$ , below which the graph contains only small connected components, and above which precisely one large component emerges (Erdős & Rényi 1960). When  $q > \frac{\ln N}{N}$ , the graph is almost surely connected (Erdős & Rényi 1960).

Profits for the complete network as a function of sharing activity are illustrated in Figure 6. Note that the complete network behaves similarly to our centralized network—profits decrease with sharing initially, until a critical point is reached and sharing activity becomes profit-increasing. The initial dip in profit is more pronounced for the inefficient (equal-pay budget balancing) mechanism, since firms are less able to appropriate consumers’ valuations.

The primary difference between the results for the complete network and those of the centralized star is that the shift to group pricing, and its accompanying rapid increase in profits, occur at a very low level of  $q$  under a complete structure (note the horizontal axis scale difference between Figure 6 and Figures 3(a) and 5(a)). This is driven by the fact that the number of possible links in the complete network is significantly higher than in the star (on the order of  $N^2$  rather than  $N$ ) and thus many more links form at a given  $q$ . Thus, the complete network quickly exhibits a large connected component. Overall, the complete network seems to suggest robustness for our model, as its insights

---

<sup>9</sup>Finding a complete subgraph of size  $k$  was one of the first problems shown to be NP-complete (Karp 1972). In particular, the solutions involve transcendental equations, a problem that plagues even many simpler models of random graphs (Newman 2003).

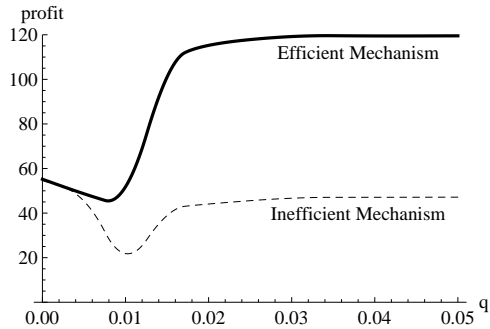


Figure 6: Profit for the complete network structure

are qualitatively similar to those of the centralized network, differing primarily in the threshold level of  $q$ .

## 6.2 Homophily

So far, we have assumed that all links within the consumer network are equiprobable. While consistent with most of the literature on random graphs, this does not capture the fact that consumer valuations and network formation may not be independent. In reality, individuals may be more likely to have social ties with those that are similar to them. The resulting correlation among the valuations of group members is termed “homophily” (Lazarsfeld & Merton 1954) or “assortativity” (Newman 2003). Homophily may arise from one of two processes. First, homophily can be a case of “similarity breeds connection” (Lazarsfeld & Merton 1954, McPherson et al. 2001, Newman 2003), where individual valuations are distributed independently, and groups tend to form among individuals with similar valuations. McPherson et al. (2001) note that this conception of homophily influences connections in a host of networks, including those formed for information exchange. Alternatively, one can envision that homophily might be a function of the opposite dynamic, “connection breeds similarity.” In this interpretation, consumers form groups according to some random process (through work relationships, friends, etc.), and then each group collectively determines a group valuation. In the specific context of unethical behavior among group members, both of these processes appear to be significant (Brass et al. 1998). While similarity breeding connection and connection breeding similarity are distinct social processes, the primary effect of either is that group members’ valuations become correlated. We remark that “similarity breeds connection”

is the traditional interpretation of homophily in the literature on social network theory that provides the methodological basis for this paper (McPherson et al. 2001). Thus, this interpretation is our focus here. However, the “connection breeds similarity” case is also addressed at the end of this section.

### 6.2.1 Similarity Breeds Connection

We examine the impact of correlated valuations within a group by introducing a link probability that depends on the difference in consumers’ valuations. In our model, a consumer is completely characterized by a valuation of the product. We use this valuation,  $v$ , to introduce assortative mixing by scalar properties Newman (2003, §III). Although a number of other consumer characteristics might drive homophily, including age, occupation, and income (Girvan & Newman 2002, Newman 2003), these will lead to similar formulations if they are correlated with valuations, as income may be.<sup>10</sup> We use the following function to transform the unbounded difference between a pair of consumer valuations,  $v_i - v_j$ , into a probability,  $q_{i,j}$ , bounded by zero and one:

$$q_{i,j} \equiv \frac{1}{1 + x|v_i - v_j|^y} \quad (4)$$

The probability of a link,  $q_{i,j}$  is decreasing in the absolute value difference. The parameter  $y$  is a mixing scalar (see Newman (2003), section III, p.4) that controls the degree of preferential attachment, while  $x$  is a scaling parameter. We investigated three different values of  $y$  (0, 0.5, and 5), to capture the effect of homophily of various strengths. For each value of  $y$ ,  $x$  was calibrated to maintain an *average* link probability of  $q$ . The results reported are for  $q = 0.5$  (results for other  $q$  do not change the insights). The no homophily case ( $y = 0$ ) is equivalent to our main model, where link probability is independent of valuations. At a high degree of homophily ( $y = 5$ ), preferential attachment is very strong and the likelihood of a link between individuals with highly disparate valuations is extremely small. This may be analogous to very strong cliques in high schools (a classic setting for homophily) where groups are very likely to consist only of

---

<sup>10</sup>Specifically, if homophily is driven by a trait correlated with valuation, then we lose no generality by concentrating on the reduced-form correlation structure among valuations.

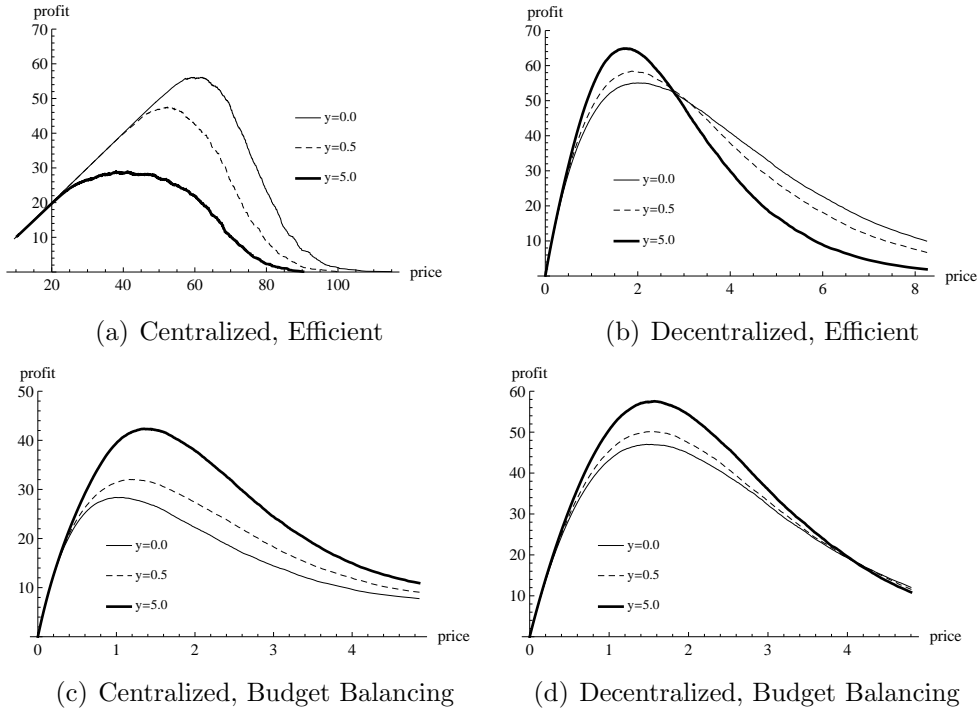


Figure 7: Profit when links form through preferential attachment under no ( $y = 0$ ), intermediate ( $y = 0.5$ ), and high ( $y = 5$ ) degrees of homophily.

sufficiently similar individuals. The intermediate homophily case ( $y = 0.5$ ) reflects a context wherein the probability of a link between two people of very different valuations is smaller than for ones with more similar valuations, but is still non-trivial.

In terms of the decision mechanisms employed by groups, there are several reasons to believe that groups formed by homophily are likely to use efficient decision mechanisms.<sup>11</sup> However, for completeness, we examine centralized and decentralized network structures using both efficient and budget balancing mechanisms. Our results are obtained by simulating 5,000 networks for each value of the mixing scalar  $y$ . Each simulation step proceeds as follows: (i) draw a value for each of 150 nodes from the exponential distribution; (ii) compute the probability of each potential link given equation (4); and (iii) randomly draw edges according to the computed probabilities. The potential links and each group's willingness to pay are computed based on the network structure and group decision mechanism, as before.

<sup>11</sup>Homophily implies some knowledge of the preferences of other group members. Thus, strategically understating one's value becomes more difficult. In addition, since homophily likely leads to stronger group ties, this implies higher levels of voluntary cooperation in contribution to a public good (VanVugt & DeCremer 1999).

The effects of homophily on firm profits are illustrated in Figure 7, which depict profit as a function of price.<sup>12</sup> While not immediately apparent from that figure, we remark that the shapes of the optimal profit curves do not change with homophily. This includes centralized network profits, which still exhibit the double-peaked shape as with our base model. However, in order to more easily observe the impact of homophily on optimal profits, in Figure 7 we “zoom in” to the peak at which profits are maximized. Although the core intuition from our main analysis persists, the presence of homophily does suggest some nuances to the insights, which we discuss below.

Intuitively, one might expect homophily to reduce firm profits, since one of the advantages of group formation is that it reshapes demand by pooling disparate consumers together, statistically averaging high and low value consumers (Bakos et al. 1999). This result is observed as expected for efficient centralized networks, where the values of group members depend on the value realized by the central node. A high-value central node leads to a group of high-value consumers, while a low-value central node leads to a group of low-value consumers. As the firm is unaware of the central node’s value, a priori, homophily merely increases the variance of demand faced by the firm.

However, we find that the insights of Bakos et al. (1999) do not hold for the other three cases illustrated in Figure 7. In the decentralized, efficient mechanism, profits increase with the level of homophily. Here, homophily has the effect of aggregating low-value consumers (of which there are many), while leaving high-value outliers as singletons. This focused aggregation among low-value consumers is ideal from the firm’s standpoint, and the effect of stronger homophily is higher profit for the firm. Furthermore, when groups use the inefficient budget balancing decision mechanism, the profit impact of homophily is unambiguously positive. With equal payment by all group members, a higher valuation by one member need not increase the group’s willingness to pay. Hence, the firm cannot appropriate this heterogeneity of valuations within groups, and it prefers group members to have similar valuations—that is, it prefers homophily. Overall, we find that, unless the firm faces a centralized consumer network which employs an efficient decision mechanism,

---

<sup>12</sup>Graphs for profit as a function of  $q$  confirm the intuition discussed below. They are omitted here for brevity but are available from the authors.

profits are enhanced when homophily is present in the sharing group formation process.

### 6.2.2 Connection Breeds Similarity

An alternative to the traditional conception of homophily is the idea that a group, instead of bringing similar people together, shapes the valuations of its randomly-grouped members. In other words, connection breeds similarity. Regardless of exactly how the group influences valuations, the valuation of a group of size  $k$  can be represented by a single random variable with a cumulative distribution function  $G_k(v)$ . Notably, in this context the group valuation  $G_k(v)$  is not a convolution of  $k$  random variables, since each group member’s valuation is entirely characterized by the valuation of the group. In other words, there are no individual valuations to mix. It is straightforward to obtain a profit equation directly analogous to the expression obtained for independent group formation in Claim 1, substituting  $G_k(v)$  for  $F_k(v)$ :  $\Pi(p) = p \sum_{k=1}^N (1 - G_k(p)) \tau_k$ . The derivation of this follows directly from the proof of Claim 1, with the substitution of  $G_k(v)$  for  $\Pr\{p \leq v_k\}$ , since the group’s valuation is no longer a  $k$ -fold convolution.<sup>13</sup> The fact that this profit function has a similar structure to the one analyzed in Claim 1 suggests that our methodology can be applied to the case where connection breeds similarity as well.

## 6.3 Ethical Considerations

In this section we consider the possibility that some consumers may never participate in sharing, perhaps for ethical or moral reasons. In the terminology of Desai et al. (2011), these consumers may have very high “psychological or moral costs” associated with engaging in illegal activities. For brevity the results provided are for the efficient group mechanism. No new insights are obtained if budget balancing within groups is assumed instead.

The analysis of a subset of consumers who never share in a centralized network is relatively straightforward. Consider a fraction of consumers,  $\alpha \geq 0$ , who never share

---

<sup>13</sup>Beyond accommodating this conception of homophily, this general formulation without restriction on  $G_k(v)$  suggests that this model structure can be adapted to several inquiries beyond those considered here. For example, one can incorporate costly sharing by specifying  $G_k(v)$  to be the  $k$ -fold convolution of individual valuations net of the costs of producing  $k - 1$  copies. Within-group network externalities can also be incorporated by specifying a  $G_k(v)$  that stochastically dominates the convolution of the valuations of  $k$  unconnected individuals.

even when connected to other consumers. Assuming that the central node does share (else, no sharing occurs for any  $q$ ), the probability of a sharing link forming is analogous to  $\hat{q} = q(1 - \alpha)$ . Thus, the probability of a group of size  $k$  forming,  $\tau_k$ , is reduced since it is now a function of  $\hat{q} < q$ . An increase in non-sharing consumers pushes this effective  $q$  lower, making the insights from our main model for smaller  $q$  more likely to be relevant. Since sharing is less beneficial for the firm at lower  $q$ , the existence of a segment of non-sharing consumers constrains the firm's price reaction for the consumers who are willing to share. If  $q$  is sufficiently high, the gradual introduction of non-sharing consumers does not change the firm's focus from pricing to sharing groups, but results in lost sales from the non-sharing consumers. In these cases, the firm would always prefer that all consumers consider sharing.

For the decentralized network, the impact of non-sharing consumers is more involved. Randomly positioning non-sharing consumers within a ring structure would create methodological complications since non-sharers would 'break' the ring into a series of linear segments. We take a simplified approach here, modeling the  $\alpha$  non-sharing consumers as singletons, with the remaining  $1 - \alpha$  consumers (who will share with connected consumers) in a ring network. The impact of non-sharing in this context depends on the optimal pricing scheme in our main model. If pricing primarily for individuals is optimal in the main model, then very little changes when we assume that  $\alpha$  consumers will certainly remain individuals. This can be seen in Figure 8(a), which shows profits as a function of price when  $q$  is low ( $q = 0.05$ ) for various  $\alpha$ . At this level of  $q$  the firm in our main model is still pricing primarily for individuals (see Figure 2 and Claim 2). The fact that three curves in Figure 8(a) are nearly perfectly overlapping shows that, as  $\alpha$  increases, neither price nor profits change dramatically in this context. Contrast this with Figure 8(b), which shows profits as a function of price when  $q = 0.95$  for various  $\alpha$ . At this level of  $q$ , the firm has shifted under our main model to pricing primarily for groups (see Figure 2 and Claim 2). Thus, when some consumers are designated as singletons by assumption, the firm faces the exact same dilemma as in the centralized network—pricing for individuals vs. pricing for groups. This results in a double-peaked

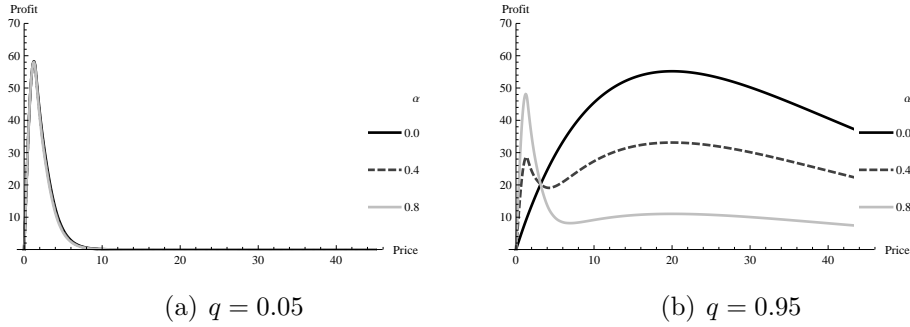


Figure 8: Price vs. profit for the decentralized network, for various non-sharing proportion  $\alpha$

profit function, representing profits from each of these two pricing schemes, as seen in Figure 8(b). As in the centralized case from our main analysis, the firm’s optimal price is determined by which of the two peaks is higher. The value of  $\alpha$  at which the profit from pricing to groups (second peak) is higher than the profit from pricing to individuals (first peak) is exactly the point at which group prices become optimal.<sup>14</sup> In this way,  $\alpha$  displays a tipping point that has interesting managerial implications. If many consumers are unwilling to share ( $\alpha$  is high), then individual pricing is optimal and the firm benefits from additional non-sharing individuals. If few consumers are unwilling to share ( $\alpha$  is low), then the firm would prefer that none exhibited this restraint, since it has already made the shift to pricing for groups. In other words, if a culture of sharing already exists, the firm might prefer that as many consumers as possible adopt that culture, while the existence of a large amount of non-sharers may influence a firm to discourage sharing among the few who do.

Overall, the results from the model extensions in this section suggest robustness of our key findings. Specifically, the fundamental effect of sharing on firm profits is quite similar across all extensions. This generality derives from the fact that, as established by Claim 1, the expected number of groups of size  $k$ ,  $\tau_k$ , and the distribution of the valuation of a group of size  $k$ ,  $F_k(v)$ , are sufficient to determine the optimal price of an information good that is shared within consumer networks. We find that, as our parameter of interest

<sup>14</sup>While up to this point we have used the terms “pricing for individuals” and “pricing for groups” rather loosely in the context of the decentralized network, this analysis allows us to be more precise. Specifically, if a firm would respond to an increase in  $\alpha$  with a discontinuous (alt. gradual) reduction in price, then it is currently pricing primarily for groups (individuals). It can be shown that this shift in price response occurs precisely where the probability of a sale to a singleton is equal to the per-consumer probability of a sale to a group.

$q$  changes, the effect on group sizes  $\tau_k$  is primarily a function of the network structure. Furthermore, the valuation of a group of a particular size  $F_k(v)$  is primarily a function of the group decision mechanism. Numerous caveats and finer-grained insights are obtained when we consider other contextual factors such as link probability (exogenous, driven by homophily, or moderated by ethical concerns). These caveats are quite interesting, and some are counterintuitive, but none represent a dramatic departure from our core insights in §4. The managerial implications of our findings are summarized and discussed in the following section.

## 7 Managerial Implications and Future Research

In this paper we use a graph theoretic approach to analyze how sharing of information goods affects monopolist profits when the social network structures that give rise to groups of consumers are explicitly considered. To our knowledge, ours is the first paper to consider probabilistic group formation via social networks in order to purchase an information good, and the pricing and profit implications of such activities under different group decision mechanisms. In particular, we analyze three types of network structures—decentralized, centralized, and complete—and two types of decision mechanisms within groups, efficient and equal-pay budget balancing. We find that both network structure and decision mechanism have important pricing and profit implications for the firm. When groups make decisions efficiently, firm profits are never reduced by decentralized sharing provided that the firm responds to sharing optimally. This optimal response is nearly always a modest price increase, although a large price hike is warranted if sharing crosses a relatively high threshold. When sharing is centralized, if the level of sharing is low then a price increase cannot offset the negative impact of sharing, and profits suffer. However, under this structure the large price hike is optimal at much lower levels of sharing, and above that point further sharing does increase firm profits. When sharing within groups is inefficient (equal-pay budget balancing), the pricing insights do not change, but the profit impact of sharing is more negative, since groups do not fully aggregate the valuations of group members.

From a managerial standpoint, our findings suggest an important link between the strength of anti-piracy efforts in a given market and a firm's optimal response to sharing activity. Government support of anti-piracy efforts varies widely, and it is minimal in some markets (Gopal & Sanders 1998). We suggest that, in a market with weaker anti-piracy laws, consumers will place less importance on concealing their illicit sharing. Following Baker & Faulkner (1993), this relative unimportance of concealment implies that sharing networks would tend to be centralized in nature. On the other hand, if anti-piracy laws are strict, then concealment becomes more important, and decentralized networks are more likely. Our results enable us to suggest a link between this logic and a firm's global pricing strategy for information goods. As shown above, a firm can be much more aggressive in shifting to a pricing strategy that targets groups when the underlying network is centralized. Furthermore, if such a group pricing strategy is optimal, then profits are always increasing with additional sharing activity. The implication, then, is that firms should take a different approach to pricing depending on the strength of anti-piracy laws. In markets with strong anti-piracy laws, sharing networks are likely to be decentralized, so a increase to group pricing is unlikely to be justified. However, the firm can benefit from sharing by enacting small-scale price increases. In markets with weak anti-piracy laws, sharing networks might be centralized, and a shift to group pricing is more likely to be optimal.

A product category perspective can also be applied to our findings. Sharing of business-related product categories, such as financial reports or real estate listing services, might involve individuals who are less inclined to reveal their true valuations. In such a context, an equal pay approach akin to our budget balancing mechanism may be more likely. Our results suggest that, in these cases, firms find it more difficult to reap benefits from sharing. Thus, managers in such product categories should be more aggressive in combating sharing. On the other hand, managers for products typically shared among friends, such as music, may be able to reap the benefits of social sharing comparatively easily, since the group mechanism is more likely to be efficient. In essence, while firms may not directly observe group mechanisms, the nature of the product may

be a good indication of them. The example of sharing music is interesting since the negative effects of sharing are often argued in the context of this category. Our results offer a different outlook, suggesting that products shared among friends are *less* likely to reduce firm profits when prices are adjusted accordingly.

Several limitations to the insights provided in this paper should be noted. First, for parsimony we assume that the individual propensity to share  $q$  is independent of price. As noted above, an explicit link between price and  $q$  would require an ad hoc assumption regarding the functional form of the relationship. However, we conjecture that the effect of such a link on our insights can be suggested without additional analysis. Consider the region of  $q$  in which, in our main model, an increase in  $q$  leads to higher profits. If  $q$  increases with  $p$ , then any increase in price in this region will further increase  $q$  (and thereby profits). In effect, in this region the firm would increase price more than what is suggested by our results. Conversely, if the firm is in a region where higher  $q$  is detrimental to profits, the logic works in the opposite direction, and the firm will increase price less than what our results suggest.

A second limitation stems from the fact that, while substantial price increases are not uncommon for institutional customers such as libraries or video rental stores, they might encounter resistance when applied to the price faced by consumers. Thus, while our model does suggest that such actions could be profitable, a word of caution is advised, since there are likely other behavioral factors to be considered. Further research could investigate the degree of social acceptance of such a policy, and the impact of this acceptance on the firm's optimal decisions when faced with sharing of its information goods.

Finally, we use a static framework this paper, using established reduced-form representations of the strategic network formation process. The explicit modeling of sequential network formation, which might require a simpler setup than the one used in this paper, could yield additional insights into the impact of social sharing on prices and profits.

## References

Andrade, N., Mowbray, M., Lima, A., Wagner, G. & Ripeanu, M. (2005), Influences on cooperation in bittorrent communities, *in* 'Proceedings of the 2005 ACM SIGCOMM

- workshop on economics of peer-to-peer systems’, pp. 111–115.
- Bagnoli, M. & McKee, M. (1991), ‘Voluntary contribution games: efficient private provision of public goods’, *Economic Inquiry* **29**(2), 351–366.
- 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.
- Bala, V. & Goyal, S. (2000), ‘A non-cooperative model of network formation’, *Econometrica* **68**(5), 1181–1231.
- Bhattacharjee, S., Gopal, R. D. & Sanders, G. L. (2003), ‘Digital music and online sharing: software piracy 2.0?’, *Communications of the ACM* **46**(7), 107–111.
- Brass, D., Butterfield, K. & Skaggs, B. (1998), ‘Relationships and unethical behavior: A social network perspective’, *The Academy of Management Review* **23**(1), 14–31.
- BSA (2008), Fourth annual BSA and IDC global software piracy study. <http://www.bsa.org/globalstudy>. Accessed 02/21/2008.
- Burton-Taylor (2009), Global financial information survey - revenue, sharing and piracy estimates. Report, Burton-Taylor International Consulting LLC.
- Cheng, H., Sims, R. & Teegen, H. (1997), ‘To purchase or pirate software: an empirical study’, *Journal of Management Information Systems* **13**(4), 49–60.
- Clarke, E. H. (1971), ‘Multipart pricing of public goods’, *Public Choice* **11**, 17–33.
- Conner, K. R. & Rumelt, R. P. (1991), ‘Software piracy: an analysis of protection strategies’, *Management Science* **37**(2), 125–139.
- Croson, R. (2007), ‘Theories of commitment, altruism and reciprocity: evidence from linear public goods games’, *Economic Inquiry* **42**(2), 199–216.
- Dawes, R. M., McTavish, J. & Shaklee, H. (1977), ‘Behavior, communication, and assumptions about other people’s behavior in a commons dilemma situation’, *Journal of Personality and Social Psychology* **35**(1), 1–11.
- Dearden, J. & Einolf, K. (2003), ‘Strategy-proof allocation of fixed costs’, *Review of Economic Design* **8**(July), 185–204.
- Deroian, F. (2009), ‘Endogenous link strength in directed communication networks’, *Mathematical Social Sciences* **57**(1), 110–116.
- Desai, P., Purohit, D. & Vernik, D. (2011), ‘Downloadable music set free: the flip side of DRM protection’, *Marketing Science* (forthcoming).
- Erdős, P. & Rényi, A. (1960), On the evolution of random graphs, *in* ‘Publication 5. Institute of Mathematics’, Hungary: Hungarian Academy of Sciences, pp. 17–61.

- Falk, A. & Kosfield, M. (2003), It's all about connections: Evidence on network formation. Working paper.
- Feri, F. (2007), 'Stochastic stability in networks with decay', *Journal of Economic Theory* **135**(1), 442–357.
- Gachter, E. F. S. (2000), 'Cooperation and punishment in public goods experiments', *American Economic Review* **90**(4), 980–994.
- Girvan, M. & Newman, M. E. J. (2002), 'Community and structure in social and biological networks', *Proceedings of the National Academy of Sciences* **99**(12), 7821–7826.
- Givon, M., Mahajan, V. & Muller, E. (1995), 'Software piracy: estimation of lost sales and the impact of software diffusion', *The Journal of Marketing* **59**(1), 29–37.
- Goeree, J., Riedl, A. & Ule, A. (2009), 'In search of stars: Network formation among heterogeneous agents', *Games and Economic Behavior* **67**(2), 445–466.
- Gopal, R. D., Bhattacharjee, S. & Sanders, G. L. (2006), 'Do artists benefit from online music sharing?', *Journal of Business* **79**(3), 1503–1533.
- Gopal, R. D. & Sanders, G. L. (1998), 'Software piracy: Key issues and impacts', *Information Systems Research* **9**(4), 380–397.
- Herman, M. (2009), FT sues Blackstone over alleged login abuse. [http://timesonline.co.uk/business/industry\\_sectors/media/article5644656.ece](http://timesonline.co.uk/business/industry_sectors/media/article5644656.ece). Accessed 12/22/2010.
- Hughes, D., Coulson, G. & Walkerdine, J. (2005), 'Free riding on Gnutella revisited: The bell tolls?', *IEEE Distributed Systems Online* **6**.
- Hwang, S., Lee, H. & Cho, S. (2009), 'Account-sharing detection through keystroke dynamics analysis', *International Journal of Electronic Commerce* **14**(2), 109–126.
- Jackson, M. O. & Wolinsky, A. (1996), 'A strategic model of social and economic networks', *Journal of Economic Theory* **71**(1), 44–74.
- Jain, S. (2008), 'Digital piracy: a competitive analysis', *Marketing Science* **28**(3), 610–626.
- 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.
- Kumar, R., Novak, J. & Tomkins, A. (2010), Structure and evolution of online social networks, in P. S. S. Yu, J. Han & C. Faloutsos, eds, 'Link mining: models, algorithms, and applications', New York: Springer, pp. 337–357.
- Lazarsfeld, P. & Merton, R. (1954), Friendship as a social process: A substantive and methodological analysis, in M. Berger, T. Abel & C. Page, eds, 'Freedom and Control in Modern Society', Princeton, NJ: Van Nostrand, pp. 18–66.

- Ledyard, J. (1995), Public goods: A survey of experimental research, *in* J. Kagel & A. Roth, eds, 'The Handbook of Experimental Economics', Princeton: Princeton University Press.
- Levati, M. V., Sutter, M. & Heijden, E. V. D. (2007), 'Leading by example in a public goods experiment with heterogeneity and incomplete information', *Journal of Conflict Resolution* **51**(5).
- Liebowitz, S. J. (1985), 'Copying and indirect appropriability: Photocopying of journals', *The Journal of Political Economy* **93**(5), 945–957.
- McPherson, M., Smith-Lovin, L. & Cook, J. M. (2001), 'Birds of a feather: Homophily in social networks', *Annual Review of Sociology* **27**(1), 415–444.
- Messmer, E. (2009), Realtors sold on software that ferrets out illicit password sharing. <http://networkworld.com/news/2009/030309-realtors-database-password.html>. Accessed 12/22/2010.
- Mortimer, J. (2007), 'Price discrimination, copyright law, and technological innovation: Evidence from the introduction of dvds', *The Quarterly Journal of Economics* **122**(3), 1307–1350.
- MPAA (2008), Motion Picture Association of America: Anti-piracy. [http://www.mpa.org/piracy\\_OptDisk.asp](http://www.mpa.org/piracy_OptDisk.asp). Accessed 02/22/2008.
- MusicAlly (2005), Napster users sharing passwords to save cash. [http://www.theregister.co.uk/2005/04/08/napster\\_password\\_sharing/](http://www.theregister.co.uk/2005/04/08/napster_password_sharing/). Accessed 12/22/2010.
- Newman, M. E. J. (2001), 'The structure of scientific collaboration networks', *Proceedings of the National Academy of Sciences* **98**(2), 404–409.
- 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.
- Newman, M. E. J., Watts, D. J. & Strogatz, S. H. (2002), 'Random graph models of social networks', *Proceedings of the National Academy of Sciences* **99**(Suppl 1), 2566–2572.
- 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.
- Ostrom, E. (2000), 'Collective action and the evolution of social norms', *Journal of Economic Perspectives* **14**(3), 137–158.
- Peitz, M. & Waelbroeck, P. (2006), 'Piracy of digital products: a critical review of the theoretical literature', *Information Economics and Policy* **18**, 449–476.
- Philadelphia Inquirer (1997), Online publishers try to end password abuse. Philadelphia Inquirer, June 1, 1997.

- Pritchard, S. (1998), Money: On-line tip sheets: the penny hasn't dropped. <http://www.independent.co.uk/news/a1151732.html>. Accessed 12/22/2010.
- Shy, O. & Thisse, J. (1999), 'A strategic approach to software protection', *Journal of economics and management strategy* **8**(2), 163–190.
- Sinha, R. K. & Mandel, N. (2008), 'Preventing digital music piracy: the carrot or the stick?', *Journal of Marketing* **72**, 1–15.
- Smith, E. & Fowler, G. A. (2011), Ganging up on internet pirates. Wall Street Journal, July 8, 2011.
- Sonnier, G., Ainslie, A. & Otter, T. (2007), 'Heterogeneity distributions of willingness-to-pay in choice models', *Quantitative Marketing and Economics* **5**(3), 313–331.
- Stubblefield, A. B. & Wallach, D. S. (2000), A security analysis of my.mp3.com and the beam-it protocol. Tech. Rep. TR-00-353.
- VanVugt, M. & DeCremer, D. (1999), 'Leadership in social dilemmas: the effects of group identification on collective actions to provide public goods', *Journal of Personality and Social Psychology* **76**(4), 587–589.
- Varian, H. R. (2000), 'Buying, sharing, and renting information goods', *Journal of Industrial Economics* **48**(4), 473–488.
- Varian, H. R. (2005), 'Copying and copyright', *Journal of Economic Perspectives* **19**(2), 121–138.
- Wilson, T. (2008), Widespread account-sharing threatens corporate security, revenues. <http://www.darkreading.com/security/client-security/212002459/index.html>. Accessed 12/22/2010.