

Contribution to P2P networks: the role of sharing behaviour in direct social networks

Tushar K. Nandi

Centre for Studies in Social Sciences, Calcutta

nandi.tushar@gmail.com

Fabrice Rochelandet

ADIS, Université Paris-Sud, Paris

fabrice.rochelandet@u-psud.fr

Abstract:

The paper studies the determinants of contribution to 'peer to peer (P2P)' networks and tests the theoretical predictions about reciprocity and altruism in the presence of non-rival goods and anonymity. Using a probit model with sample selection and survey data about sharing behaviour of more than 2000 individuals in P2P and three direct social networks (family, friends and workplace), the paper tests the consistency of sharing behaviour over direct and indirect networks. Using observational data, this paper finds results that are consistent with two findings of recent experimental studies. First, our results reveal a pattern of replication of sharing behaviour from direct network to virtual networks. Second, the contribution in P2P networks is poorly determined by the factors posited in the utilitarian approach. The findings suggest that, in presence of social embeddedness of sharing behaviour, it may be difficult to design an effective legal instrument to fight illegal sharing.

Keywords:

P2P networks; social networks; copyrights; contribution; sharing; reciprocity; free riding.

1. Introduction

Recently there has been a spate of debate in developed countries about the illegal copying of copyright contents. The internet mediated file swapping communities, commonly known as peer to peer (P2P) networks, have often been the focus of this debate. The academic as well as the policy making community seems to be divided on the issue of targeting the P2P communities in the fight against illegal copying. There are at least two major reasons why P2P networks occupy the centre stage in this debate. First, P2P network is viewed as an innovation that enables designing new business models based on this transfer protocol. Second, it has the potential to be instrumental in the fight against illegal file-sharing of copyright works. In both cases, it is crucial to understand why people contribute resources for the benefit of other participants of P2P networks. In the first case, leading users to contribute more resources could enable promoters of P2P solutions to support the expansion of the network and enhance its performance¹. In the last case, governments and copyright owners might seek to lead people to contribute less and less digital contents until the disappearance of P2P sharing networks for lack of utility. Most OECD countries are trying to design legal and technical devices to eliminate or control unauthorised P2P sharing and to deter the unlicensed posting of copyright contents on user generated content (UGC) platforms.

Our paper aims at evaluating the theoretical predictions about reciprocity and free riding in the presence of non-rival goods and anonymity. If, on one hand, motivations for downloading are quite well explored by empirical studies, on the other hand, there is much less written about why people actually contribute. Hence little is known why individuals keep on contributing to the commons in the presence of massive free-riding and when this behaviour proves costly for them. In other words, why do online sharing communities thrive in spite of the theoretical predictions and empirical observations of low levels of contribution?

We investigate empirical regularities on the illegal sharing of copyright contents. The originality of our study is that we use data that contains information about individual behaviours - contributing, free riding, and non-participation - from a large heterogeneous sample in two kinds of social networks: P2P networks and direct face-to-face social networks, namely family, friends, and workplace. We empirically test hypothesis derived from utilitarian as well as social interaction perspective. We find that the contribution to P2P network is not influenced by rational self-interested motives. Rather, it is affected by norms gathered in direct social networks. In particular, our results suggest a pattern of replication of sharing behaviour from direct networks to P2P network, hence social embeddedness of sharing behaviour.

Rest of the paper is organized as follows. Next section provides a brief survey of the literature that seeks to explain the sharing behaviour and highlight the social determinants of contribution behaviour over P2P networks. The third section presents the data used in this paper. The fourth section describes the econometric model. Section five presents the estimation results. Section six concludes and envisages some policy and business implications.

2. Contribution vs. free riding in sharing networks

The empirical literature on P2P networks has often focused on the prevalence of massive free riding in P2P networks. Two issues received considerable attention: the impact of downloading on sales (Oberholzer and Strumpf, 2007; Liebowitz, 2006) and the motivations for downloading (Holm, 2003; Rochelandet and Le Guel, 2005). Though the contribution behaviour remains the key factor determining the survival of P2P networks, a very few studies has investigated the motivation for contribution behaviour (Ripeanu et al., 2006; Nandi and Rochelandet, 2008). Most of the research in this area attempts to explain the sustainability of P2P networks on the basis of predictions derived from some game theoretical framework.

In this paper, we explore the motivation for contributing copyright goods in P2P networks. From a conceptual point of view, two approaches can be adopted in order to explain the motivation for contribution to P2P networks. These approaches – the utilitarian approach and the social interaction approach - have been used in different disciplines of social science in order to explain the contribution behaviour in the case of public good. Sociological notions such as altruism, reciprocity and other-regarding self-interest have been as instrumental as the notion of rationality in these approaches. Nandi and Rochelandet (2008) adopt a utilitarian approach in their empirical investigation of the motivations for contribution in P2P networks. They find that P2P contribution is poorly determined by rational self-interested motives. This paper extends Nandi and Rochelandet (2008) by incorporating a perspective derived from social psychology along side the utilitarian approach. In particular we explore how social environment and social norms affect the contribution behaviour in P2P networks. In the next subsection we review the basic building blocks of this approach and how they can help explain the contribution behaviour in P2P networks.

The influence of social environment

It proves difficult to explain the contribution to P2P content-sharing networks only as the action of self-interested rational individuals. Social norms constitute another way to explain generalized reciprocity towards (anonymous) others. Many factors can produce and foster the influence of social norms on individual motivations: education that induces norms of conduct, imitation, judgement of others and in particular, 'reliable authority' (one's social neighbours for instance), ex-ante positive feelings towards others, sense of equity, etc. Individuals cooperate according to their social preferences and status: normative hedonism, charismatic codes. The question is how they acquire these norms of behaviour and how this process of acquisition in turn influences their sharing behaviour.

P2P sharing networks constitute a specific type of computer-mediated communities. As in most social networks, two features can influence the willingness to contribute of their members: interconnections between them inside the virtual community and their own social life. We focus here on these two possible explanations. The first one is suggested by Strahilevitz (2003) according to which motivations for contribution are determined by 'charismatic codes' that are built by the promoters or the organizers of the P2P communities. The second one suggests that the motivations for contribution behaviour are determined in many aspects of social life, in particular, in other social networks with similar objective. Sharing behaviours are both forged and a key element of socialization in direct social networks (friends, family and workplace). Social norms built in these close-knit networks are replicated over P2P networks where the main object is also to swap copies of copyright goods.

Charismatic codes

Concerning the first explanation, P2P sharing networks work on the basis of rules. They can be technical rules which are imposed by the provider of sharing technologies in order to compel users to cooperate: automatic rating, downloading conditional on authorized uploading, and so on. Hence, technical rules can structure the networks by providing extrinsic motivation². Alternatively, it can be private norms, namely 'charismatic codes' that are assimilated to intrinsic motivation. It is precisely what Strahilevitz (2003) refers in his social psychological analysis. Sharing occurs in network because of the intentional perception given to network users that contribution is a widespread and normal practice in the P2P community³. *"Members of the loose-knit file-swapping networks cooperate with each other largely because the networks' creators give their users a distorted picture of the community ... [N]orms of reciprocity ... are internally enforced - through file-swappers' desire to avoid*

feelings of guilt and selfishness or to experience the warm glow associated with group solidarity."⁴

In the same way, Golle et al. (2001) highlight the communication services offered by Napster (chat-rooms, newsletter, direct messaging), which were aimed at fostering a "sense of community" among its users and therefore, to induce them to contribute and reciprocate.

Norms from the real world

We suggest another approach to explain the role of behavioural norms in shaping the intrinsic motivations for contribution in P2P content sharing networks. We argue that contributors can acquire the sense of altruism and reciprocity outside the virtual communities. We consider how the social life of P2P users matters in shaping behaviour in virtual communities. Most of the time, individuals conform to conventions and replicate them in all social aspects of their life, including in the virtual communities. Our conception is in contrast to evolutionist approaches that explain the emergence and persistence of altruistic behaviours. Most of these approaches assume that altruistic population survive due to the retaliation against those who do not respect the cooperative norm to the detriment of the utility of altruist agents (Bowles and Gintis, 2004).

To illustrate our proposition, we can consider the notion of 'docility' as defined by Simon (1990, 1993). Bounded rationality prevents optimal behaviour for achieving fitness of human beings. They have neither complete and certain knowledge, nor the computational skills to determine the exact consequences of their choices. Hence, 'docility' contributes to the fitness of human beings. It is defined as "the tendency to depend on suggestions, recommendations, persuasion, and information obtained through social channels as a major basis for choice". More than one's own experience, information obtained from others and social interactions influence individual choice. Altruism and the sense of reciprocity of sharers reflect in certain aspects their docility on P2P networks. Higher a user's acquired sense of reciprocation and altruism when trying to increase her fitness through social interaction in physical social networks, the more likely she is to share in P2P networks. To some extent, altruism, generalized reciprocity and community interest created by ongoing interaction of the members of these close-knit networks create contribution motivations for certain users of P2P networks.

Replication behaviour from personal to anonymous contexts is becoming a major issue in behavioural and experimental economics. For example, results obtained in a dictator game

setting often exhibit positive donations (Camerer, 2005). To explain the existence of this kind of non-selfish behaviours among anonymous players, several authors have developed interpretative models around concepts such as “inequality aversion” (Bolton and Ockenfels, 1998; Fehr and Schmidt, 1999), “altruism” (Andreoni and Miller, 2002), “egocentrism” (Cox et al., 2002) and Rawlsian “social welfare” preferences (Charness and Rabin, 2002). However, these models have often been criticised on the ground that observed behaviours in labs seem in large part disconnected from everyday life observations on anonymous donations. This kind of behaviour would be rather significant into familiar interactions (with family, friends) than in strictly anonymous contexts (see Bardsley, 2008, for a survey).

Recent works by Charness and colleagues (Charness et al., 2007; Charness and Gneezy, 2008) contribute crucially to the debate concerning the role of emotional proximity in these kinds of experimental settings. These authors, finding that in previous experimentation “many field interactions are conducted with neither complete anonymity nor complete familiarity” (Charness and Gneezy, 2008, p. 30), extend classical works by Hoffman et al. (1996) and Bohnet and Frey (1999) on “social distance” and “anonymity”. Charness and Gneezy (2008) test the effect of social distance on players’ behaviour in developing two experimental procedures applied on dictator and ultimatum game settings. In the first one, individuals follow standard dictator and ultimatum rules; in the second one, each participant is informed of the name of their opponents. They conclude that (i) the name revelation has its main generosity effect in the dictator game settings, (ii) this additional information has no significant effect on sharing in the ultimatum game, where the possibility of credible sanction exists.

Charness et al. (2007) use a lost-wallet game played in classrooms as well as in the Internet to measure differentials in positive reciprocity behaviours. Their paper is highly informative for our own study by the fact that the aim of the authors is to apprehend other-regarding behaviours in a “delocalized interactions among strangers” context. They find that interactions through the Internet reveal an important amount of reciprocity concerns, even if the intensity of reciprocity is decreasing with social distance. The authors mention: “we were surprised at how little difference we observed between the treatments, particularly since our classroom experiments are nearly the opposite with respect to social distance. Reciprocity appears to be a factor, even in a virtual experiment. To the extent that reciprocity could be a feature of virtual international business perhaps cooperative behaviour is sustainable” (Charness et al., 2007, p. 101).

To summarize, reciprocal behaviours seem to be slightly affected by social distance, while altruistic behaviours are more. To our knowledge, our study is one of the first to test such proposals on a large scale, combining the familiar and anonymous dimension of social interactions. Moreover, we test the regularities regarding cooperation and social distance found in experimental settings with data collected from survey.

3. Data and variables

Our analysis is based on primary data collected in January and February 2005 about sharing behaviour in P2P and three direct social networks (family, friends and workplace). A total of 2533 individuals were surveyed using a paper survey and a Web-based survey in France. To simplify missing data correction, we chose to use the list-wise deletion approach (Allison, 2001). The sample bias due to the Web-based survey has been corrected using a post-stratification method implemented with an SAS software macro named CALMAR and developed by the French national institute for statistics and economic studies (INSEE)⁵.

Our survey collects information about the individual participation in sharing of copyright digital contents in P2P network and in three direct networks – family, friends, and workplace. It also provides information about contribution or free riding behaviour of individuals upon participation in each network. Table 1 presents the definition and descriptive statistics of the variables constructed from the information contained in our data set and used for the empirical analysis of this paper. The explanatory variables correspond to the conceptual framework developed in the previous section.

After omitting for the missing values, a sample of 2062 individuals is used for the analysis. Table 1 shows that the sample consists of 51% individuals who participate in P2P network. Of these participants, 44% contributes new contents to the P2P network and 56% free rides. The participation rate (% who share digital contents) is little higher than 50% in two direct networks – family and friends. However, it is quite low (33%) in case of workplace network. Looking at the contribution vs. free riding behaviour among participants of the direct networks, we observe that around 70% are contributors and around 30% free riders. Our sample consists of 19% female and 81% male individuals. The sample distribution of age, education, occupation and household income shows that there is considerable variation in data in terms of socio-demographics characteristics.

In addition to the influence of sharing behaviour in direct network on the same in P2P network we postulate that the latter can be affected by experience gathered through social interaction. The variable ‘Herding’ is introduced to capture the effect of interaction with

copiers in the individual's social neighbourhood. The question is to what extent the number of copiers in the social neighbourhood of an individual (which he can observe and/or with whom he can communicate and share experiences) influences positively his cooperative behaviour over P2P networks. The underlying assumption is that P2P users acquire cooperative routines in their direct social networks.

Table 1: Definition and summary statistics of variables

Variable	Definition	Mean
<i>P2P</i>		
Participation	equals 1 if the individual participates in P2P sharing network, 0 otherwise.	0.51
Contribution	equals 1 if the individual contributes in P2P network, 0 otherwise.	0.44
Free-ride	equals 1 if the individual receives contents from P2P network but does not provide, 0 otherwise.	0.56
<i>Family</i>		
Participation	equals 1 if the individual participates in sharing contents in family, 0 otherwise.	0.52
Contribution	equals 1 if the individual contributes in sharing contents in family, 0 otherwise.	0.72
Free-ride	equals 1 if the individual receives contents from family members but does not provide, 0 otherwise.	0.28
<i>Friends</i>		
Participation	equals 1 if the individual participates in sharing contents with friends, 0 otherwise.	0.57
Contribution	equals 1 if the individual contributes in sharing contents with friends, 0 otherwise.	0.70
Free-ride	equals 1 if the individual receives contents from friends but does not provide, 0 otherwise.	0.30
<i>Workplace</i>		
Participation	equals 1 if the individual participates in sharing contents in workplace, 0 otherwise.	0.33
Contribution	equals 1 if the individual contributes in sharing contents in workplace, 0 otherwise.	0.68
Free-ride	equals 1 if the individual receives contents from workplace but does not provide, 0 otherwise.	0.32
<i>Gender</i>		
Female	equals 1 if female, 0 otherwise.	0.19
<i>Age</i>		
age <= 24 yrs	equals 1 if the individual is less than 25 years old, 0 otherwise.	0.14
25 yrs <= age <= 40 yrs	equals 1 if the individual is aged between 25 and 40 years, 0 otherwise.	0.51
age > 40 yrs	equals 1 if the individual is more than 40 years old, 0 otherwise.	0.35
<i>Education</i>		
less than BAC	less than BAC/BAC Pro	0.10
BAC/BAC Pro	BAC/BAC Pro (high school graduate, business, technical)	0.16
BAC+1+2	BAC+1+2 - some college (not 4 yrs degree).	0.27
BAC+3+4	BAC+3+4 - BS or more.	0.21
more than BAC+5	more than BAC+5 - MA.	0.26
<i>Occupation</i>		
Freelance	Freelance executive.	0.36
Intermediate occupations	Intermediate occupations, skilled worker, worker.	0.38
Retired	Retired.	0.06
Student	Student.	0.07
Unemployed	Unemployed.	0.13
<i>Monthly household income</i>		
less than 1000 euro	less than 1000 euro per month.	0.11
b/w 1000 and 1500 euro	between 1000 and 1500 euro per month.	0.17
b/w 1500 and 2000 euro	between 1500 and 2000 euro per month.	0.17
b/w 2000 and 2500 euro	between 2000 and 2500 euro per month.	0.15
b/w 2500 and 3000 euro	between 2500 and 3000 euro per month.	0.14
b/w 3000 and 3500 euro	between 3000 and 3500 euro per month.	0.08
b/w 3500 and 4000 euro	between 3500 and 4000 euro per month.	0.07
b/w 4000 and 5000 euro	between 4000 and 5000 euro per month.	0.06
more than 5000 euro	more than 5000 euro per month.	0.05
<i>Herding (number of copiers in social neighbourhood)</i>		
none	none.	0.11
b/w 1 to 5	between 1 to 5 persons.	0.26
b/w 6 to 15	between 6 to 15 persons.	0.21
more than 15	more than 15 persons.	0.41
Cultural diversity	Do you think the legal market for music does not offer enough variety? - 1 for agree and completely agree, 0 otherwise.	0.39
<i>Experience with internet</i>		
less than 1 year	less than 1 year.	0.12
b/w 1 to 2 years	between 1 to 2 years.	0.18
b/w 2 to 3 years	between 2 to 3 years.	0.28
more than 3 years	more than 3 years.	0.42
Willingness to pay *	willingness to pay for a P2P network that gives unlimited access (in Euro).	6.09 (6.69)
Legal risk *	perceived legal risk of being caught for using P2P network: 0 no risk, 1 low, 2 medium and 3 high.	1.65 (0.90)
Technical risk *	perceived technical risk of being infected by virus and other malicious contents while using P2P network: 0 no risk, 1 low 2 medium and 3 high.	1.60 (1.02)
Ethics *	An index of ethical concern based on the opinion on four aspects: copying (i) threatens the existence of the market for music and CD (ii) threatens the income of artists and others involved (iii) does not respect the work of the artists and others involved, and (iv) is bad in general. Coding: 1 do not agree, 2 somehow agree, 3 agree, and 4 strongly agree.	6.11 (2.31)
Broad-band connection	equals 1 if the individual has broad-band connection at home, 0 otherwise.	0.83

Number of observations: 2062

* standard deviation in brackets.

A dummy variable representing quest for cultural diversity is also incorporated as an additional motivation for sharing in P2P network. We have also information about the individual's internet experience. It is possible that more experienced internet users are likely to participate more in virtual communities, and be aware of costs and benefits of participation.

The variable 'Willingness to pay' represents the sum that the individual would accept to pay to have an unlimited access to digital contents through a P2P network. According to the utilitarian approach this variable is positively correlated with contribution behaviour over P2P network. The underlying hypothesis is that the more individuals value a sharing network, the more they derive utility from its existence (and persistence), and hence the more they are willing to contribute to feed it.

The last group of variables corresponds to the cost factors in a utilitarian approach. They are legal, technical risks, ethical concern and the quality of internet connection. The variables representing the risks refer to the perceived risks associated with unauthorised sharing, namely the perceived likelihood of being caught and sanctioned or infected by undesirable contents like virus and other malicious contents. They are supposed to be negatively associated with the participation in P2P network. The underlying assumption is that the greater the perceived risk of using P2P file-sharing networks, the greater is the perceived cost associated with participation. The variable 'Ethics' stands for an index that sums up the ethical concerns of the individual regarding sharing of copyrighted works. It indicates the psychological 'costs' the individuals bear when they feel ethically wrong while sharing. This variable is supposed to impact negatively any use of P2P file-sharing networks. The last variable stands for the broad-band internet connection. It is likely that individuals who have broad-band connection at home are more likely to participate in P2P networks since the chance of experiencing network congestion is much lower.

Table 2: Cross tabulation of sharing behaviour in P2P by sharing behaviour in direct networks

		P2P			
		Non-participation	Participation	Contributiomm	Free-ride
Family	Non-participation	53.31	46.69	40.13	59.87
	Participation	44.17	55.83	46.80	53.20
	Contribution	38.13	61.87	51.36	48.64
	Free-ride	60.07	39.93	28.21	71.79
Friends	Non-participation	53.96	46.04	39.71	60.29
	Participation	44.46	55.54	46.52	53.48
	Contribution	38.73	61.27	51.40	48.60
	Free-ride	55.06	44.94	37.14	62.86
Workplace	Non-participation	51.19	48.81	40.77	59.23
	Participation	43.26	56.74	49.35	50.65
	Contribution	40.13	59.87	54.95	45.05
	Free-ride	51.00	49.00	40.03	59.97

Table 2 gives cross tabulation of sharing behaviour in P2P network by sharing behaviour in other networks. Each block of numbers (percentages) gives the break down for a particular direct network. In each row, the first two numbers give the break down into non-participation and participation in P2P network. And the last two numbers give further break down of the P2P participants into contributor and free rider.

The 1st row in the 1st block of numbers show that 53% of those who do not exchange contents among family members do not participate in P2P network. Among the rest (47% who participate in P2P network), 40% are contributor and 60% free rider in P2P network. The 2nd row shows that 56% of sharers in family network participate in P2P network, and 47% of these P2P participants are contributor. The last two rows of the block give further detail of the family network participants. They show that more than half of the contributor (free rider) in family network are participant (non-participant) in P2P network. Conditional on participation in P2P network more than half of the contributor (free rider) in family network remain contributor (free rider) in P2P network. The patter is very similar for other direct networks (other blocks of numbers in Table 2).

The general pattern that emerges from this table is that those who do not exchange contents in direct networks are less likely to participate in P2P network and, upon participation, more likely to be free rider in P2P. The contributors in direct networks are more likely to participate in P2P network than the free riders in direct networks. The contributors (free riders) in direct networks are more likely to be contributors (free riders) in P2P network. This descriptive analysis points to a pattern of behavioural replication from direct network to

virtual network. The econometric analysis that follows in next sections provides further evidence in this regard.

4. Estimation strategy

For each network (direct and indirect), we observe an individual's choice regarding participation/non-participation in sharing digital contents. We also observe whether an individual contributes or free rides upon participation to a network. The aim of the empirical analysis that follows is to estimate the effect of sharing behaviour in direct network on the contribution behaviour in P2P network. The contribution in a network is defined as providing new contents to the network irrespective of whether the individual receives contents from other members of the network⁶. This definition excludes the case where an individual obtain contents from a network and makes them available to other members of the same network. We define this latter behaviour as free riding since it does not feed the network with new contents which are vital for the survival of a sharing network.

Since we observe sharing behaviour in P2P network of the individuals those who participate in P2P network, we use a sample selection model. The selection equation explains the participation of individuals in P2P network as opposed to non-participation. For the participants we define the outcome equation that explains contribution as opposed to free riding behaviour in P2P network. Formally the model can be defines as follows.

Let y_1^* define a latent variable for the propensity to participate in P2P network. Suppose that y_1 is a dummy variable that takes value 1 for participation in P2P network, 0 otherwise. The observable variable y_1 is related to y_1^* as follows.

$$\begin{aligned} y_1^* &= x_1' \beta_1 + u_1 \\ y_1 &= 1 \text{ if } y_1^* > 0 \\ &= 0 \text{ otherwise} \end{aligned}$$

where x_1 is the vector of explanatory variable for participation in P2P network, β_1 the vector of parameters, and u_1 a random term.

For the participants of P2P network, let y_2^* define a latent variable for the propensity to contribute in P2P network. Suppose that y_2 is a dummy variable that takes value 1 for contribution in P2P network, 0 otherwise. It is important to note that y_2 is observed only when $y_1 = 1$. The observable variable y_2 is related to y_2^* as follows.

$$\begin{aligned}
y_2^* &= x_2' \beta_2 + u_2 \\
y_2 &= 1 \text{ if } y_2^* > 0 \\
&= 0 \text{ otherwise}
\end{aligned}$$

where x_2 is the vector of explanatory variable for contribution behaviour in P2P network, β_2 the vector of parameters, and u_2 a random term. Assuming normal distribution for the error terms,

$$\begin{aligned}
u_1 &\sim N(0,1) \\
u_2 &\sim N(0,1) \\
\text{Corr}(u_1, u_2) &= \rho
\end{aligned}$$

we have a probit model with sample selection. The model parameters (β_1, β_2 and ρ) are estimated by maximum likelihood method where the likelihood function incorporates the probability of participation and the probability of contribution conditional on participation in a sharing network.

5. Estimation results

In this section we present our estimation results. The probit model with sample selection is estimated by maximum likelihood method. We incorporate variables representing sharing behaviours of one direct network in a single specification⁷. Before turning to the results a few comments about the model specification is warranted. We introduce the variables related to cost of using P2P networks only in selection equation, not in P2P contribution equation. The reason is that these variables – legal and technical risks, ethics and broad-band connection – are more likely to affect the participation decision to the P2P network. They are, at the same time, less likely to be distinguishing factors for contribution vs. free ride decision in P2P network. The variable concerning the value of P2P network – willingness to pay – is included only in P2P contribution equation, not in selection equation. The idea is that higher values attributed to P2P network are likely to enhance contribution in order to ensure persistent future stream of benefits from P2P network. We include the direct network participation variable in the selection equation in order to examine whether participation in direct network is associated with participation in virtual network. However, detailed break down of participation in direct network – into contribution and free riding – is incorporated in P2P contribution equation. This allows us to examine the patten of replication of sharing behaviour from direct to virtual network. The socio-demographic and internet

experience variables are included in both the equations. Other specifications of the model are also used as robustness check of our main results (more below).

Table 3: Estimates from the probit model with sample selection

	(i)				(ii)				(iii)			
	P2P participation		P2P contribution		P2P participation		P2P contribution		P2P participation		P2P contribution	
	Coefficient	S.E	Coefficient	S.E	Coefficient	S.E	Coefficient	S.E	Coefficient	S.E	Coefficient	S.E
<i>Family network</i>												
Participation	0.153 ***	0.060										
Contribution			0.164 *	0.085								
Free-riding			-0.352 ***	0.130								
<i>Friends' network</i>												
Participation					0.093	0.062						
Contribution							0.179 **	0.088				
Free-riding							-0.315 ***	0.120				
<i>Workplace network</i>												
Participation									0.085	0.065		
Contribution											0.253 ***	0.093
Free-riding											-0.182	0.125
<i>Gender</i>												
Female	-0.258 ***	0.082	0.220 *	0.115	-0.247 ***	0.081	0.209 *	0.115	-0.236 ***	0.081	0.195 *	0.114
<i>Age</i>												
25 yrs <= age <= 40 yrs	-0.188 *	0.109	0.092	0.129	-0.182 *	0.110	0.082	0.129	-0.196 *	0.110	0.100	0.129
age > 40 yrs	-0.439 ***	0.124	0.258 *	0.153	-0.430 ***	0.124	0.225	0.154	-0.446 ***	0.124	0.218	0.155
<i>Education</i>												
BAC/BAC Pro	-0.140	0.120	-0.076	0.146	-0.137	0.120	-0.103	0.147	-0.136	0.120	-0.102	0.146
BAC+1+2	-0.339 ***	0.111	-0.227	0.143	-0.339 ***	0.111	-0.259 *	0.144	-0.338 ***	0.111	-0.246 *	0.143
BAC+3+4	-0.407 ***	0.118	-0.113	0.153	-0.409 ***	0.118	-0.150	0.154	-0.408 ***	0.118	-0.131	0.154
more than BAC+5	-0.498 ***	0.121	0.087	0.156	-0.502 ***	0.121	0.072	0.157	-0.498 ***	0.121	0.084	0.156
<i>Occupation</i>												
Freelance	0.004	0.105	-0.043	0.139	0.003	0.105	-0.044	0.140	-0.005	0.105	-0.084	0.139
Intermediate occupations	0.009	0.098	-0.094	0.130	0.017	0.098	-0.092	0.132	0.011	0.099	-0.142	0.131
Retired	0.143	0.152	-0.334	0.214	0.154	0.152	-0.276	0.215	0.156	0.152	-0.307	0.214
Student	0.062	0.167	-0.028	0.199	0.074	0.167	-0.058	0.201	0.083	0.167	-0.016	0.200
<i>Monthly household income</i>												
b/w 1000 and 1500 euro	0.074	0.118	0.197	0.148	0.072	0.118	0.222	0.149	0.067	0.118	0.229	0.148
b/w 1500 and 2000 euro	-0.045	0.121	0.131	0.152	-0.038	0.121	0.151	0.153	-0.045	0.121	0.151	0.152
b/w 2000 and 2500 euro	-0.044	0.126	-0.042	0.160	-0.043	0.126	-0.023	0.161	-0.048	0.126	0.000	0.161
b/w 2500 and 3000 euro	0.052	0.130	0.184	0.163	0.057	0.130	0.225	0.164	0.049	0.130	0.235	0.164
b/w 3000 and 3500 euro	0.136	0.147	-0.270	0.186	0.142	0.147	-0.223	0.187	0.138	0.147	-0.215	0.186
b/w 3500 and 4000 euro	0.041	0.155	0.000	0.195	0.045	0.155	0.016	0.197	0.040	0.155	0.025	0.196
b/w 4000 and 5000 euro	-0.063	0.163	-0.012	0.213	-0.058	0.163	0.006	0.213	-0.073	0.163	0.052	0.212
more than 5000 euro	-0.004	0.173	0.281	0.226	0.005	0.173	0.259	0.226	-0.003	0.173	0.267	0.225
<i>Herding</i>												
b/w 1 to 5	0.215 **	0.108	-0.161	0.161	0.217 **	0.108	-0.164	0.162	0.219 **	0.108	-0.122	0.162
b/w 6 to 15	0.500 ***	0.112	-0.095	0.172	0.499 ***	0.112	-0.097	0.173	0.504 ***	0.112	-0.041	0.173
more than 15	0.525 ***	0.106	-0.150	0.168	0.519 ***	0.106	-0.158	0.169	0.525 ***	0.106	-0.099	0.169
Cultural diversity	0.310 ***	0.061	0.032	0.087	0.311 ***	0.061	0.025	0.086	0.314 ***	0.061	0.029	0.086
<i>Experience with internet</i>												
b/w 1 to 2 years	-0.128	0.110	0.188	0.154	-0.124	0.110	0.203	0.155	-0.120	0.110	0.182	0.155
b/w 2 to 3 years	0.116	0.104	0.168	0.141	0.120	0.104	0.177	0.142	0.126	0.104	0.185	0.142
more than 3 years	0.012	0.105	0.252 *	0.142	0.021	0.105	0.277 **	0.143	0.023	0.105	0.257 *	0.142
Willingness to pay			-0.009	0.006			-0.010 *	0.006			-0.010 *	0.006
Legal risk	0.022	0.032			0.024	0.032			0.023	0.032		
Technical risk	-0.007	0.029			-0.008	0.029			-0.009	0.029		
Ethics	-0.077 ***	0.013			-0.078 ***	0.013			-0.078 ***	0.013		
Broad-band connection	0.738 ***	0.081			0.737 ***	0.081			0.737 ***	0.081		
Constant	-0.145	0.237	0.099	0.311	-0.128	0.238	0.093	0.311	-0.087	0.236	0.058	0.309
Number of observations	2062		1060		2062		1060		2062		1060	
Log likelihood				-1951.120				-1952.790				-1955.870
Rho				-0.566				-0.550				-0.550

Note: *, ** and *** stand for significance at 10%, 5% and 1% level, respectively.

Table 3 gives the estimation results from probit model with sample selection. Three specifications are used for three direct networks. In specification (i), family network variables are incorporated. The first two columns give the estimates and standard errors for selection equation. The last two columns give the same for P2P contribution equation. The specifications (ii) and (iii) give the estimation results that incorporate friends and workplace network variables, respectively.

The first thing to note in the table is that the results are very similar for three specifications. The participation in direct network is positively associated with the participation in P2P network, though the coefficient is significant only for family network participation. Higher age groups are less likely to participate in P2P network, as are the higher educational groups. Occupation and household income appear to be insignificant for P2P participation decision. Higher the number of copier in social neighbourhood higher is the likelihood of P2P participation. Quest for cultural diversity positively affects participation in P2P network. Surprisingly, legal and technical risks appear to be insignificant determinant of P2P participation. However, ethics negatively and possession of broad-band connection positively affect P2P participation.

Turning to the estimation results for P2P contribution equation, we find that socio-demographics factors are not significant determinants of P2P contribution. Female participants are more likely to be contributors, though the estimates are marginally significant in all specifications. Herding and cultural diversity, though important for P2P participation, are insignificant for P2P contribution. There is slight evidence that higher internet experience is positively associated with contribution in P2P network. This result could confirm that adherence to the social norms associated with the Web is fostered as people use Internet. Surprisingly enough, the coefficient of willingness to pay is negative. However, the evidence is too weak and marginally significant only in two specifications. If anything, the negative coefficient of willingness to pay indicates that people who are willing to pay for P2P network consider the network as way to purchase digital contents.

The contribution in direct network has a positive effect on the contribution in P2P network. The effect is the strongest for workplace network followed by friends and family networks. Similarly, free riding in direct network has a negative effect on P2P contribution. It is strongest for family network, followed by friends network. In case of workplace network, this effect is not significant.

The last row of the table gives the estimate of the correlation (ρ) between error terms of two equations. The higher value of the correlation implies that the error terms are

highly correlated – an evidence in support of the choice of selection model. The negative value of the correlation coefficient implies that as far as the unobservable factors are concerned they affect P2P participation and P2P contribution decisions in opposite directions. This is not surprising given the massive free riding in P2P network.

We perform a set of robustness checks for our main results using different specifications⁸. In the estimation above we introduce willingness to pay in the contribution equation arguing that those who value P2P network are likely to contribute in order to ensure persistent future stream of benefits from P2P network. However, it is possible that users get value from P2P network by downloading digital contents, not by uploading and they do not perceive the free riding problem. In order to examine this possibility we re-estimate the model incorporating willingness to pay in participation equation. The estimate of willingness to pay remains insignificant even in participation equation and other results do not change.

The costs associated with participation in P2P networks, in particular the perceived legal risk, are incorporated in the participation equation. However, the legal risk can be important for P2P contribution if people perceive that copyright infringements can only be enforced when a user redistributes the content, not when she downloads it. In other words, people who fear legal action should be reluctant to upload, not to download content in P2P networks. We re-estimate the model with legal risk variable in the P2P contribution equation. The estimate of the variable remains insignificant and other results do not change. It is possible to think that ethical concern affects both P2P participation and P2P contribution. As another robustness check we incorporate ethical concern in both participation and contribution equations. The variable does not exert a significant effect on contribution. However, as before it remains negative and significant for P2P participation.

6. Conclusion

Explaining why people contribute to P2P networks is crucial to the understanding how these sharing networks actually work and, above all, the conditions of their (non)viability. Indeed, the availability of files and diversity on a P2P network depend strictly on the willingness of some peers to upload copies. Our econometric results suggest that the motivations for contributing are poorly determined by rational self-interested behaviour. Rather, the behaviour in P2P networks appears to be a replication of the behaviour on direct social networks. Sharing behaviours are embedded into local interactions structures.

Current strategies to fight against illegal P2P file-sharing try to influence their utility. One strategy consists in reducing their absolute utility. In fact, it amounts to increasing the

costs associated with contribution in order to dry up the supply of contents over P2P networks. But our findings suggest that this strategy might be inefficient. Our results are sharply in contrast to the assertion of Krishnan et al. (2004) that "increasing the cost of sharing can reduce the number of sharers and above a certain point lead to network collapse [...] Increasing the implicit cost of sharing is by increasing the legal risks to individual network users from sharing copyrighted information."⁹ In fact, our findings suggest the opposite: Copyright enforcement - in particular, increased sanctions, legal suits against individual copiers - has no impact on the contribution behaviour. Contributors get into the habit of sharing contents with their direct social 'neighbours' and then they replicate the behaviour mainly in accordance to these habits. So, if nobody stops sharing in her neighbouring or is subject to legal prosecution, the copier will keep on copying and sharing contents online. That might explain the failure of current enforcement strategies.

In particular, our findings cast some doubts about the efficiency of the recent bill of the French government aiming at putting an end to unauthorized P2P file-sharing¹⁰. The 'réponse graduée' consists in, firstly, sending a warning by email to the copyright infringers who upload content on P2P networks; then - if the P2P user keep on uploading - sending a warning by registered mail; and finally - if the infringer continues uploading - imposing a fine on her. At best, this new form of copyright enforcement could eliminate the current P2P networks used to share contents. But, even if it occurs, our results suggest that any new sharing technology that would be candidate for replacing P2P technologies might be automatically fuelled by the 'supply' of contents contributors due to the very existence of already-existing practices.

That might explain the failure of current enforcement strategies (increased sanctions, legal suits against individual copiers)¹¹. Our results are consistent with the result of Bhattacharjee et al. (2006) that "even after these legal threats and the resulting lowered levels of file sharing, the availability of music files on these networks remains substantial." It seems difficult for the law to create the conditions for free-riding inside P2P networks if, in fact, uploading decisions are strongly determined by social norms that are built outside the P2P networks - when individuals directly shares copies in physical social networks. The viability and sustainability of virtual content-sharing communities depend on the traditional sharing practices. Perhaps, copyright industries could succeed in the sort run to technically close the current P2P networks by leading more people to free-ride *on these networks*. But it seems very difficult to prevent indefinitely the online replication of existing and common direct sharing.

Instead, copyright industries can rather build innovative business models - increasing the quality of services and diversity of contents available in the legal markets - to compete efficiently with P2P content-sharing by decreasing the relative utility of P2P networks. These business models should even try to extract the value from these sharing communities (Le Guel and Rochelandet, 2006). And copyright law should stimulate innovation by implementing compulsory licences in order to facilitate the acquisition of copyrights and increase the size of catalogue supplied to consumers on legal markets.

Though our study is able to shed light on the social embeddedness of sharing behaviour in P2P network, caution is required in interpreting the results as causality. The results can well mean that people come to P2P networks with a predisposition towards certain behaviours. Given the general difficulty in delineating causal effect in social interaction and the scarcity of empirical works on the interaction among different social networks, we believe that our study makes valuable contribution to enhance our understanding in a rather little researched area.

¹ Content producers and online retailers envisage the viability of digital business models using sharing technologies to legally distribute copyright contents (music, movies, software etc.) to end users. For instance, the main advantage of P2P architecture is to lower the costs of distribution, computation, storage, and so on. Also, user generated content (UGC) websites - such as YouTube and Dailymotion - enable the distribution of diversified video contents without incurring the costs of producing them. The common feature of these two types of P2P networks is that the shared contents are voluntary contributed by individual users.

² We draw the classical distinction between extrinsic and intrinsic motivations for contributing. The former is external to individuals - for instance, monetary rewards, sanctions, social capital or recognition from family or friends, whereas the latter is embedded in the person herself (the motivation that comes from inside, her psychology, beliefs...). See Kandell and Lazear (1992) for a parallel distinction between "shame" and "guilt."

³ The study of Strahilevitz (2003) considers only the legal aspect of P2P file-sharing and not the economic impact (postulated as negative) or the economic interest of P2P networks as potential business models.

⁴ The acceptance of charismatic codes is facilitated by the fact that file-sharing cannot be assimilated to a large-scale for-profit piracy. Instead, it was promoted initially as being an altruistic sharing. This message was fostered by media and public opinion polls showing that people supported P2P sharing and were against legal prosecutions against file-sharers. In this specific context, promoters of P2P technology exploited the norm of reciprocity by magnify the cooperative behaviours and discouraging the non-cooperative ones. Thus, "*when an individual receives a benefit that obviously results from the cooperation of others, she internalizes a feeling of indebtedness The best way to remove these feelings of guilt is for her to reciprocate directly.*"

⁵ For information and download the CALMAR Macro, see http://www.insee.fr/en/home/home_page.asp / 'Classification, Definitions — Methods' page, 'Statistical Tools'.

⁶ Note that this definition includes reciprocal behaviour.

⁷ Since multicollinearity among the variables for the sharing behaviour in different direct networks can confound the estimates, we estimate the model using only one direct network variables at a time. However, an additional estimation, not reported here, shows that our results do not change qualitatively when sharing behaviours of all direct networks are included in a single specification.

⁸ For the sake of space, the results of robustness checks are not reported in the paper. However, they can be obtained from the authors.

⁹ In the same way, Asvanund et al. (2004) put that "*The more recent strategy adopted by copyright holders of bringing legal action against violators may be more successful even though the proportion of users who are targeted is a small fraction of the total number of users. The success of this strategy depends on raising the implicit cost of sharing for users by raising their legal risks. Increased sharing costs will then raise their propensity to free-ride and may ultimately reduce the utility offered by "illicit" file trading over P2P networks enough to make the legitimate purchase of the music an attractive option for users.*" Or Strahilevitz (2003) states

that *"If my account is correct, it suggests that the copyright industries' efforts to control copyright infringement on peer-to-peer networks have been wrongheaded. Rather than moving sequentially against the various post-Napster networks, the copyright industries might have adopted various strategies to create a norm of free-riding, thereby cutting off the cooperative uploading on which these networks rely."*

¹⁰ Recently a similar law has been enacted in Sweden.

¹¹ We show elsewhere (Rochelandet and Le Guel, 2005) that P2P sharing is only a small part of copying behavior: copiers mainly share with their social 'neighbours'. Copiers decide whether or not to share mostly according to the decisions of their social relations. Two consequences: (1) if nobody in her neighbouring stops sharing or is subject to legal prosecution, the copier will keep on sharing contents ; (2) Everyone can switch from a kind of sharing network to another if she perceives a legal risk. Such an embeddedness of virtual networks facilitates the diffusion of innovation in sharing technologies.

Acknowledgement

We thank the members of ADIS research group, in particular Fabrice Le Guel and Didier Lebert, as well as the participants of the 2008 SERCI Congress for helpful discussions. This work was supported by the French National Agency of Research, research program ANR-05-JCJC-0204-01.

References

- Allison, P.A.: 2001, 'Missing Data', *Sage University Papers Series on Quantitative Applications in the Social Sciences*, 07-136, Thousand Oaks, CA: Sage.
- Andreoni, J. and J. Miller: 2002, 'Giving according to GARP: an experimental test of the consistency of preferences for altruism', *Econometrica* 70(2), 737-753.
- Asvanund, A., K. Clay, R. Krishnan and M. Smith: 2004, 'An empirical analysis of network externalities in peer-to-peer music sharing networks', *Information Systems Research* 15(2), 155-174.
- Bardsley, N.: 2008, 'Dictator game giving: altruism or artefact?', *Experimental Economics* 11(2), 122-133.
- Bhattacharjee, S., R.D. Gopal, K. Lertwachara and J.R. Marsden: 2006, 'Impact of legal threats on online music sharing activity: an analysis of music industry legal actions', *Journal of Law and Economics* 49(1), 91-114.
- Bohnet, I. and B.S. Frey: 1999, 'Social distance and other-regarding behavior in dictator games: comment', *American Economic Review* 89(1), 335-339.
- Bolton, G. and A.Ockenfels: 1998, 'Strategy and equity: an ERC-analysis of the Güth-van Damme game', *Journal of Mathematical Psychology* 62(2-3), 215-226.
- Bowles, S. and H. Gintis: 2004, 'The evolution of strong reciprocity: cooperation in heterogeneous populations', *Theoretical Population Biology* 65(1), 17-28.
- Camerer, C.F.: 2005, 'Behavioral economics', *World Congress of the Econometric Society*, London 18-24 August.
- Charness, G. and U. Gneezy: 2008, 'What's in a name? Anonymity and social distance in dictator and ultimatum games', *Journal of Economic Behavior & Organization* 68(1), 29-35.
- Charness, G., E. Haruvy and D. Sonsino: 2007, 'Social distance and reciprocity: an Internet experiment', *Journal of Economic Behavior & Organization* 63(1), 88-103.
- Charness, G. and M. Rabin: 2002, 'Understanding social preferences with simple tests', *Quarterly Journal of Economics* 117(3), 817-869.

- Cox, J., K. Sadiraj and V. Sadiraj: 2002, 'Trust, fear, reciprocity and altruism', Working Paper. University of Arizona.
- Fehr, E. and K.M. Schmidt: 1999, 'A theory of fairness, competition and cooperation', *Quarterly Journal of Economics* 114(3), 817-868.
- Golle, P., K. Leyton-Brown, I. Mironov and M. Lillibridge: 2001, 'Incentives for Sharing in Peer-to-Peer Networks', Working paper, Stanford University.
- Hoffman, E., K. McCabe and V.L. Smith: 1996, 'Social distance and other-regarding behavior in dictator games', *American Economic Review* 86(3), 653-660.
- Holm, H.J.: 2003, 'Can Economic Theory Explain Piracy Behavior?', *Topics in Economic Analysis & Policy* 3(1), 1-15.
- Kandel, E. and E.P. Lazear: 1992, 'Peer pressure and partnership', *Journal of Political Economy* 100 (4), 801-817.
- Krishnan, R., M.D. Smith, Z. Tang and R. Telang: 2004, 'The virtual commons: Why free-riding can be tolerated in file sharing networks', Working paper, Carnegie- Mellon University.
- Le Guel, F. and F. Rochelandet: 2006, 'The willingness to pay for online music in the presence of copying: An empirical investigation', ADIS working paper, Université de Paris-Sud.
- Liebowitz, S.J.: 2006, 'File-Sharing: Creative Destruction or just Plain Destruction?', *Journal of Law and Economics* 49(1), 1-28.
- Nandi, T. K. and F. Rochelandet: 2008, 'The incentive for contributing digital contents over P2P networks: An empirical investigation', *Review of Economic Research on Copyright Issues* 5(2), 19-35.
- Oberholzer, F. and K. Strumpf: 2007, 'The Effect of File Sharing on Record Sales. An Empirical Analysis', *Journal of Political Economy* 115(1), 1-42.
- Ripeanu, M., M. Mowbray, N. Andrade and A. Lima: 2006, 'Gifting Technologies: A BitTorrent Case Study', *First Monday* 11(11), November.
- Rochelandet, F. and F. Le Guel: 2005, 'P2P music sharing networks: Why the legal fight against copiers may be inefficient', *Review of Economic Research on Copyright Issues* 2(2), 69-82.
- Simon, H.A.: 1993, 'Altruism and Economics', *American Economic Review* 83(2), 156-161.
- Simon, H.A.: 1990, 'A mechanism for social selection and successful altruism', *Science* 250(4988), 1665-1668.

Strahilevitz, L.J.: 2003, 'Charismatic Code, Social Norms, and the Emergence of Cooperation on the File-Swapping Networks', *Virginia Law Review* 89(3), 505-595.