A Theory of Voluntary Pooled Public Knowledge Goods and Coalition Formation

Tom Dedeurwaerdere*and Paolo Melindi Ghidi†

Status: working paper presented at ISNIE 2013, Florence, 22nd of June 2013

Abstract

In this paper we develop a theoretical model of the mechanisms behind the voluntary provision of impure public goods in coalitions in presence of important social networks effects. The model builds on the large empirical literature on coalitions for voluntary provision of pooled public knowledge goods, such as in social networks of open source software developers and consortia producing open data repositories. This literature shows that, under some conditions, the provision of public goods can be facilitated by social network effects such as group identity and social approval of individual pro-social attitudes. To integrate these effects in standard public good theory this paper follows a two-step strategy, based on the introduction of two types of impureness in standard public good theory: (1) impureness related to private excludable benefits (so-called ancillary private benefits of the public good); (2) impureness related to the satisfaction of the individuals social preferences. In a first step, the paper analyses the introduction of combined public and private benefits in coalition theory with standard preferences. In a second step the model is broadened to the case of impureness related to the social preferences. The analysis shows that, when the private benefit component of the impure public good is important, the effect of the social preferences on the coalition formation is ambiguous: with increasing/decreasing relative weight of the social approval of individual pro-social attitudes compared to the relative weight of the social group identity, the coalition size to be reached will be respectively larger/smaller compared to the coalitions formed by agents with standard preferences. Applications of the theoretical model to large-scale surveys of Free/Libre/Open-Source (FLOSS) software developers confirm the results of the model.

Key Words: Coalition formation; social dilemma; group belonging; knowledge commons

JEL Codes: H42 - Publicly Provided Private Goods; D71 Associations

*Université Catholique de Louvain and Fonds National de la Recherche Scientifique (FSR-FNRS)

†Université Catholique de Louvain

1 Introduction

The last two decades have been characterised by an unprecedented and rapid advance in technological innovation and the proliferation of digital information and communication networks (Giddens, 1999). This process has significantly altered the ways in which various types of knowledge, such as scientific information, research materials and traditional knowledge, can be produced, distributed, purchased and shared on the global scale. In particular, direct public availability of research results and innovations has proven to provide major socio-economic benefits. As a result, publicly available knowledge goods produced in social networks have emerged in almost all scholarly disciplines and knowledge contexts (Benkler 2006). In particular, the empirical literature shows that publicly available knowledge goods can be provided spontaneously through the formation of coalitions implemented in private ordering arrangements (Reichman and Uhlir 2003; Hess and Ostrom, 2007).

To understand the emergence of these public knowledge pools, it is possible to rely on the economic theory of public goods. The latter has developed a broad set of models that take into account ancillary private monetary and non-monetary benefits for individuals that contribute to the public good provision. However, some questions still remain unanswered in this literature. First, what is the role of these private benefits and social preferences in motivating coalition members to join a public knowledge pool? In particular, what is the danger of such a privatising approach to the building of coalitions, especially in situations where the market price for equivalent amounts of private goods is substantially lower than the price paid, or time invested, for receiving ancillary benefits by contributing to the networked knowledge good? Second, how to explain the results of a vast empirical literature, which shows that the effects of social preferences on facilitating the process of coalition formation are often ambiguous? Indeed, social group identity effects and reputational effects in social networks can work in the opposite direction. In particular, these effects can make coalition formation more difficult if the group shows a strong disapproval for receiving substantial ancillary private pay-off when contributing to a public good. To answer these questions, this paper develops a theoretical model of the mechanisms behind the provision of public knowledge goods through

coalitions, with the view to analyse the ambiguous effect of social preferences on the coalition formation. To the best of our knowledge, this paper is one of the first attempts to build a theory to understand these ambiguous effects in the context of coalition theory.

To build the model of coalitions with social preferences, this paper proceeds in two steps: in a first step the paper extends the theory of coalition formation with public goods to the case of impure public goods with standard preferences. In a second step, it broadens this benchmark model to account for social preferences, based on a social psychology approach to the role of social preferences in group behaviour.

The first step builds upon the public good model of Kosfeld (Kosfeld et al. 2009). Kosfeld applies the model of coalition formation, initially developed in context of the negotiation of International Environmental Agreements (IEA) (Carraro and Siniscalco 1993, Barret 1994), to the problem of pure public good provision. To extend the model of Kosfeld to the case of impure public goods, this paper draws upon the general theory of impure public goods with standard preferences proposed by Cornes and Sandler (1984; 1994) and further developed by Kotchen (2006, 2007, 2009) and Vicary (1997, 2000)[1].

The second step builds upon the vast social psychological literature that analysis the role of social preferences in extra-role behaviour and organizational citizenship behaviour. As shown in this literature, two main social psychological dimensions play a role in behaviours in which short-term sacrifice leads to long-term collective organisational benefits: group belonging (organisational commitment) and altruistic motives (such as pro-social motives; civic virtues; etc.) (Lepine et al., 2002). Depending on the context, decline in one of these dimensions has shown to lead to a more or less steep decline in the provision of collective benefits in organisations (Joireman et al., 2006; Penner and Finkelstein, 1998). The paper will apply these insights to the particular case of organisations created through coalitions of individual

[1]One of the contributions of the work of Cornes and Sandler is to show that free-riding over other agents contribution to public goods decreases when goods give ancillary private benefits to the agents. A stylised example given by Sandler of such an impure public good production model is the case of international development cooperation, where in some arrangements donor countries also derive private benefits from such cooperation, such as through selling of technology, in addition to the global public good benefits related to the overall increase in economic development and poverty alleviation (Sandler and Arce 2007).

agents who pool knowledge assets through private ordering arrangements.

The main results of the paper are the following. The comparative statics shows that, in the model with standard preferences, the effect of the private ancillary benefits on the optimal coalition size is negative. This means that coalition formation will be easier, as it is sufficient to reach a smaller coalitions size. Adding social preferences for group belonging and individual pro-social behaviour introduces a second effect on the coalition size. When the private benefit component of the impure public good is sufficiently large, the effect of social preferences on coalition formation is ambiguous: with increasing/decreasing relative weight for individual pro-social reputation compared to social group identity, the coalition size to be reached will be respectively larger/smaller compared to the coalitions formed by agents with standard preferences. Finally, an application of the theory to a heterogeneous coalition composed of two groups with different levels of social preferences shows an alternative strategy for overcoming social dilemmas in groups with homogeneous social preferences and small ancillary benefits, which is the broadening of the homogeneous group to a larger heterogeneous group. Applications of the theoretical model to large-scale surveys of Free/Libre/Open-Source (FLOSS) software developers confirm the results of the model.

The paper is organised as follows. Section 2 provides the basic motivation and facts of our analysis of voluntary contribution to publicly available knowledge goods and presents some stylised facts. Section 3 theoretically analyses the analytical conditions under which social dilemmas arise and how these can be overcome through coalition formation in a model with standard preferences. Section 4 introduces social preferences and the concept of social network coalitions. In Section 5 we present an application of the model assuming heterogeneous preferences. Section 6 concludes.

2 Motivation and Facts

A vast body of empirical research has shown the effectiveness of initiatives for the provision of public knowledge goods through voluntary mechanisms (Lessig 2001, Benkler 2006, Hess and Ostrom 2007, David 2008). Two features of successful economic and institutional arrangements for providing public knowledge goods high-

lighted in this literature are especially relevant for this paper: first, the importance of reaching an adequate group size for the knowledge generation process to proceed and, second, the role of private ancillary benefits on the willingness of agents to contribute to the public good coalition.

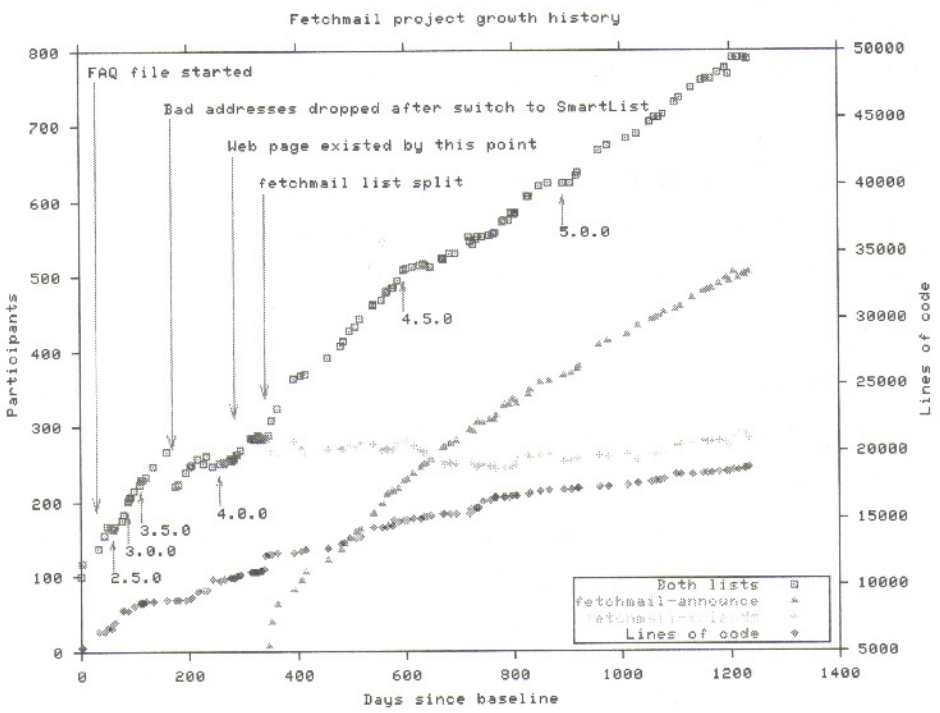
The first feature has been widely analysed in empirical surveys of open source software communities. For example in a recent large-scale survey, Schweick (2012) has shown that reaching a certain group size is a condition for developers to effectively pool their efforts and contributions. In other terms, only if the aggregate level of jointly produced public good is high enough, agents will be motivated to voluntary join the pool.

Further, when the minimal coalition size is reached, this size will theoretically also be the optimal size as there is no economic incentive for other agents to join the contributors group as they can free ride upon the public good once it is produced. This duality between the group of user-developers on the one hand and the broader user group on the other is a well-known pattern that is observed for instance in open source development projects. Indeed, in these projects, the group of users-developers grows till the group-size is sufficient for the task at hand (which various from small to larger groups), while the broader user group of the software tool continues to grow beyond that size.

To illustrate this feature we have represented in figure 1 the data of an early open source project, which is the development of the fetchmail email client, one of the first desktop email clients, which was developed to replace web mail email consultation only. The statistics in figure 1 show the data of user statistics directly pulled out from the website, from October 1996 to late 2000. The list of users that are not developers (square boxes) and the list of users that are also developers (+ signs) was split only at the release of the 4.3.0. version (at around 280 days). Remarkably, from that period the developers population remained fairly stable around an average of about 250 participants, while the user population continues a rather consistent linear growth until reaching 800 in the year 2000. This graph aptly illustrates empirically this feature of coalition theory: the knowledge pool (the source code and the subsequent improvements to the code) will be stable only if the aggregated public knowledge that the group members generate gives them a sufficient return in terms

of the problem solving performance of the publicly available software tool. However, once the adequate group size of users-developers is reached, there is no economic incentive for others to join the coalition of users-developers and the remaining users that stay outside the coalition can free ride upon the software improvements without contributing to them.

Figure 1: Development of the open source fetchmail email client



Source: Raymond, 1999, p. 215.

The second feature, which is also illustrated by the case of the fetchmail client, is that voluntary provided public knowledge goods often are impure public goods with a joint public/private character. This means that these goods generate both aggregate public and ancillary personal private benefits to the contributors, such as higher citations for researchers through increased visibility, or access to new personal competences by joining a group with high-level expertise. As a consequence, agents contributing to public knowledge pools are both driven by the public good benefits and these ancillary private benefits. Evidence for this mixed public/private character has been given in many studies of public knowledge goods in fields beyond the case of the software developers communities. One can think for example of open access

databases with tailor made data management tools that benefit specific communities (David, 2005) or hybrid funding arrangements for openly available culture products on the Internet (Lessig, 2008).

As has been shown elsewhere, the mixed public/private character can lead to contrasted effects on public good production. In particular, if the private ancillary benefit is a substitute to an equivalent private benefit that can be otherwise obtained through the purchasing of a private good on the market, the impure public good character might lead to a decrease in the willingness to contribute to the public good in situations where the market price of the substitute is sufficiently low (Cornes and Sandler, 1984). The extent to which this ambiguous effect is attenuated or exacerbated by coalition formation is one of the questions that will be addressed in this paper.

The benchmark model with standard preferences developed in the 3rd section of the paper accounts for these two basic features. Section 4 and 5 then broadens this benchmark model to develop a model which takes into account the role of social preferences in coalition formation. For doing so the analyses relies on the general economic approach of intrinsic motivations and applies this general approach to the specific case of social preferences related motivations in group dynamics.

In the general economic literature, many studies have shown that individuals can be motivated both by extrinsic motivations (motivations by external rewards such as contingent monetary rewards or rewards from a principal to an agent) or by intrinsic motivations (based on the individuals intrinsic desire to perform the task for its own sake), which leads to contrasted effects on behaviour (Benabou and Tirole, 2003). Similarly, in the specific case of public good provision, it has been shown that private contribution to a public good can be motivated by the benefit derived from the aggregate public good (the economic reward) or by altruistic motivations (the intrinsic motivations) (Andreoni, 1988).

To apply these general insights to the specific case where the intrinsic preferences relate to involvement in group processes, we rely on the well-established empirical literature in social psychology on the role of intrinsic motivations in group behaviour. This literature shows that two types of social preferences play a key contribution in the overcoming of social dilemmas that occur in the provision of public goods. First,

studies provide compelling evidence that the longing for a positive social group identity is a key determinant of engagement in group behaviour. Social psychological experiments have shown that social group identity is even in many cases the most important explanatory factor to account for various types of group behaviour related motivations such as procedural justice, fairness and supervisor ratings (Blader and Tyler, 2009). Importantly, this research shows that the effect of social group identity depends on group size: social group identity only contributes to overcoming social dilemmas in sufficiently large groups (De Cremer and Leonardelli, 2003). The second type of social preference that plays an important role in group engagement is the social approval for individual pro-social attitudes and behaviour (Lepine et al., 2002). These individual pro-social reputational effects have also been studied extensively in the context of the overcoming of social dilemmas (Suurmond et al., 2004; Bolton et al., 2005).

In the case of impure public goods, the combination of these two effects, the positive social group identity procured by the membership of a group that provides a public good and the social approval for individual pro-social behaviour, can provide contrasted effects on coalition formation. Indeed, as we will see below, if the private ancillary benefits of the public good are very substantial, the social approval is likely to decrease as the behaviour will be perceived as being based more on self-interest than on altruistic motives. Therefore, for knowledge goods with important private ancillary benefits, social preferences can also inhibit coalition formation, instead of promoting the public good coalitions. In public good theory this will be expressed by the need to reach a larger coalition size (case of difficult coalition formation), compared to coalitions which can already be stable with smaller groups of contributors.

The contrasted effects of social preferences on the size of coalitions in coalition formation processes can be illustrated by an in depth analysis of the motivations, personal attributes and behavioural patterns among free/libre and open-source (FLOSS) developers, based on the FLOSS-US 2003 survey (David and Shapiro, 2008). This web-based survey generated a wealth of data on motivations and reasons for developers to begin to work for FLOSS. Using this data, the study by David and Shapiro classifies the respondents according to their distinct motivational profiles

Table 1: Motivational Clusters among Open Source Software Developers

Cluster	Profile	Key characteristics
1	Professionals	Non-ideological, expert, self-employed or company-sponsored to collaborate on FLOSS projects
2	Aspiring hackers	No need to modify existing code but like fixing bugs and learning new programs
3	Social learners	Become better programmers, learn how programs work, work with like-minded, "give back to community," support FLOSS ideology
4	Social programmers	Experienced, employment related needs to use, modify existing code and fix bugs; project choice influenced by social connections with other developers
5	"User-innovators"	Modifying existing software unimportant, learning and interacting with like-minded others unimportant; wanted to "give back to community," and launched own project

(a) Key Characteristics of Motivational Clusters (source : David and Shapiro, 2008 : p. 384.

Cluster		Small project and large project populatons only		
		Small (1–2)	Large (>29)	Total
1 (Professionals)	%	**5.2**	**5.5**	**5.3**
	N	22	10	32
2 (Aspiring hackers)	%	**7.6**	**16.7**	**10.3**
	N	32	30	62
3 (Social learners)	%	**49.1**	**45.0**	**47.8**
	N	207	81	288
4 (Social programmers)	%	**14.0**	**12.8**	**13.6**
	N	59	23	82
5 (User-innovators)	%	**24.2**	**20.0**	**22.9**
	N	102	36	138
Total	%	**100.0**	**100.0**	**100.0**
	N	422	180	602
Pearson chi-squared(4)			11.09	
Prob > chi-squared			***0.03***	
Chi-squared goodness-of-fit(4)			60.75	
Prob > chi-squared			***0.00***	

(b) Distribution of small and large project participants by motivation profiles identified by cluster analysis of FLOSS-US survey respondent (source: David and Shapiro 2008, table 12, p. 394.

by hierarchical cluster analysis. In addition, whenever possible, the respondents in each cluster are also matched to projects of known membership sizes, revealing that the fractions of respondents from each motivational cluster for the large and the very small project ranges are statistically different, as shown in table 1.a and 1.b. below. As can be seen from this study, the three clusters of social learners, social programmers and user-innovators are present both in the small and large ranges of the project sizes, indicating that these groups have a comparatively facility to form coalitions, compared with other clusters that are only present in the large ranges. This is consistent with the fact that group belonging and social group identity related to collective good provision fosters cooperation in situations of social dilemma discussed above. In contrast, the cluster of aspiring hackers, which is composed of individualist, materially motivated programmers, which take part to FLOSS in the interest of a future career (David and Shaprio, 2008, p. 383), is by far more present in the large-size groups than in the small-size groups. This is consistent with the fact that, in presence of large private returns of the collective goods, and in the absence of strong social group identity related social preferences, coalition formation is comparatively more difficult, as the willingness to join a coalition will depend to a larger extent on the marginal return on the aggregate public good that one only obtains in larger groups.

3 Voluntary Contribution to Public Knowledge Goods: Theory with Standard Preferences

This section theoretically analyses the analytical conditions under which social dilemmas arise in the provision of impure public knowledge goods and how these can be overcome through coalition formation in a model with standard preferences. The first step of the theoretical analysis builds upon the emerging literature on coalition formation for voluntary contributions to pure public goods. To extend these models of pure public goods to the case of impure public goods, this section draws more specifically upon the general theory of impure public goods with standard preferences developed by Cornes and Sandler (1984; 1994).

It is well established in the International Environmental Agreements (IEA) liter-

ature that the formation of coalitions can partially overcome the free-rider problem in the provision of public goods. While there exist a large theoretical literature on endogenous coalition formation among countries in the context of global environmental problems (Carraro and Siniscalco 1993, Barret 1994), only few papers consider coalition formation based on the voluntary contribution to public goods by the pooling of endowments held by private actors (Kosfeld et al. 2009, McEvoy 2010).

Following these papers we consider a non-cooperative Nash equilibrium situation where agents face a social dilemma and their dominant strategy is to invest their full endowment in the private market good, i.e. the numeraire. In this situation, coalition formation is modelled through a two stage game. During the first stage, which is called the membership stage, agents decide whether or not to participate in the coalition. In the second stage, which is called the contribution stage, agents act as joint payoff maximisers and they determine how much of their endowment they will invest in the public good, knowing that they will receive a shared public benefit from the aggregated public good. The equilibrium level of participation in the coalition is obtained by solving the game using backwards induction, assuming that agents seek to maximise their expected payoffs.

We consider an economy populated by i = 1, ..., N individuals investing a part of their endowment in a certain type of knowledge good (software, data, research tools), each of whom is endowed with exogenous wealth w to be allocated between consumption of a standard private good x_i, the numeraire, and their investment in the impure knowledge good g_i with price $p > 0$, such that $x_i + pg_i \leq w_i$. For the purpose of the analysis we use a monetary endowment and a price p, but in general the endowment can also be an endowment of available time and the price the cost of the time invested in contributing to the knowledge good.

To apply coalition theory to the case of impure public knowledge goods, we consider that the knowledge good g_i is an impure public good which jointly generates $\alpha \in [0, 1]$ units that are enjoyed privately, and $\beta \in [0, 1]$ units of a public good. As in Kosfeld et al (2009), we restrict ourselves to the linear case of a sum of goods and define the payoff function as in Cornes and Sandler (1984, 1994)[2].

[2]In this simple linear case the two goods are substitutes because indifference curves are linear and

Agent i's material payoff is defined by the function $\pi_i = x_i + \alpha g_i + \delta\beta G$, where $G = \sum_{i=1}^{N} g_i$ and $\delta \in]0,1[$models the constant marginal benefit or payoff from producing the public good. This analytical formulation of the game leads to knife-edge solutions characterised by full or null purchasing of the impure public good and, therefore, the contribution or not to the public knowledge good. To remain within the non-cooperative equilibrium scenario, we assume that the technology of the impure public good is such that the price is sufficiently high compared to the payoff from the ancillary benefit $\alpha < p$[3].

The interaction of the agents generates a n-players prisoner dilemma. We assume a social dilemma scenario, which is guaranteed when $\frac{\bar{\delta}}{N} < \delta < \bar{\delta}$ where[4] $\bar{\delta} = \frac{p-\alpha}{\beta}$. This assumption implies that the strategy profile $g_i = 0$, $\forall i = 1, .., N$ is the unique non-cooperative Nash equilibrium of the game. Indeed, the game implies that efficient Pareto optimal outcomes involve a full purchasing of the impure public good $g_i = \frac{w}{p}$ because the public component of this particular good plays the same role as a public good in standard theory. Thus, players have an incentive to cooperate and voluntary pool the impure public good in order to overcome the social dilemma situation[5].

Following the two stage strategy of general coalition theory, it is now possible to compare the strategy of agents that do not participate in the coalition, to those that participate in the coalition. First, during the contribution stage it is straightforward to observe that the dominant strategy for each agent not participating in the coalition, labelled with nc, is to purchase only the pure private good. To demonstrate this, first consider the payoff of individuals outside the coalition that respectively

total utility depends on a linear combination of the quantities of each good consumed. Following Cornes and Sandler (1984), this situation is likely to increase the level of free riding in the public good game without coalitions.

[3]If the price of purchasing the impure public good is too low, i.e. $p < \alpha$ then the numeraire x_i is never acquired. In this particular case the optimal strategy for all individuals would be to demand only the impure public good because its price is lower than the marginal benefit given by the private component of the impure public good.

[4]$\bar{\delta}$ can be interpreted as marginal per capita return cut-off.

[5]As our model considers the case of knowledge goods that are available to all, the standard public good game is a particular case of our model when $\alpha = 0$ and $\beta = 1$. Public good production is given by the sum of the non-excludable component of the impure public good of individual i, βg_i and all other agents in the economy, βG_{-i}.

acquire the impure public good to those that do not acquire the impure public good:

$$\pi_{g,nc} = w_i - pg_i + \alpha g_i + \delta\beta(g_i + G_{-i}) \tag{1}$$

$$\pi_{0,nc} = w_i + \delta\beta(G_{-i}) \tag{2}$$

Let define $\Delta_{nc} \equiv \pi_{g,nc} - \pi_{0,nc}$. After some algebraical manipulations we obtain:

$$\Delta\pi_{nc} = g_i[\delta\beta + \alpha - p] \tag{3}$$

The optimal individual behaviour depends on the sign of the term in brackets in (3). It is straightforward to observe that for each agent outside of the coalition Δ_{nc} is maximised when $g_i = 0$ if $\delta < \bar{\delta}$, while when $g_i = \frac{w_i}{p}$ if $\delta > \bar{\delta}$. Under social dilemma, the latter condition can never hold. Therefore each member not participating in the coalition will never invest resources to acquire the impure public good.

Next, we consider the payoffs of members of the coalition acquiring or not the impure good. This payoff will be respectively given by:

$$\pi_{g,c} = w_i - pg_i + \alpha g_i + \delta\beta(G_c + G_{nc}) - \theta g_i \tag{4}$$

$$\pi_{0,c} = w_i + \delta\beta(G_{nc}) \tag{5}$$

with βG_c the total public good provided by the coalition via the contribution of the coalition members to the impure public good g_i and βG_{nc} the public good level produced outside the coalition. Defining with s the size of the coalition we can define $G_c = sg_i$.

As in Kosfled, we introduce a cost of participating in the coalition. In the two-stage game, the cost of participating in the coalition can be supposed to be a constant fraction of the total impure public good acquired by the agents. Therefore, in presence of identical individuals, the average participation costs coincide with individual cost, θg_i.

After some algebraical manipulations we derive $\Delta\pi_c \equiv \pi_{g,c} - \pi_{0,c}$:

$$\Delta\pi_c = g_i[\delta\beta s + \alpha - p - \theta] \tag{6}$$

The sign of the expression in the brackets in (6) is conditional on the size of the coalition and on the cost of the coalition formation. Given linearity of the payoff, members will invest their total wealth in impure public good if this sign is positive, otherwise they will invest their endowment only in the pure private commodity (the numeraire).

Turning next to the membership stage, each agent has to decide if he joins the coalition or not. Solving equation (6) allows to determine the minimum coalition size that is sufficient so that agents will join the coalition. We define with s^{min} this minimum profitable coalition size. The Nash equilibrium associated to the decision to stay or not in the coalition follows from considering the well-known stability conditions developed in DAspremont et al. (1984), largely used in the international environmental agreement literature (Carraro and Siniscalco, 1993, Barret, 1994, Ulph, 2004). More precisely, a coalition is stable if and only if it is both internally stable (no member is better off leaving the coalition) and externally stable (no agent outside of the coalition wants to join it).

Proposition 1 *If individuals are identical, there exists a coalition equilibrium of size s^* defined as the smallest integer such that*

$$s^* \geq s^{min} \equiv \frac{p - \alpha + \theta}{\beta\delta}$$

The sub-game perfect Nash equilibrium of the voluntary coalition game is given by $s^* = s^{min}$. When $s^{min} \geq N$ then a grand coalition with $s^* = N$ forms.

Proof. *Proposition 1 follows from the observation that the only non-trivial internally and externally stable coalition is the smallest profitable coalition of size $s^{min} \equiv \frac{p-\alpha+\theta}{\beta\delta}$. Knowing that under social dilemma agents outside the coalition will never purchase the impure knowledge good, we observe:*

$$\pi_c(s^{min}) \equiv w - (p - \alpha)g_i + \delta\beta s^{min} g_i - \theta g_i \geq w \equiv \pi_{nc}(s^{min} - 1)$$

if and only if $s \geq s^{min} \equiv \frac{p-\alpha+\theta}{\beta\delta}$ However, if the coalition has a size $s > s^{min}$, then at least one member has an incentive to defect while the remaining members would continue to stay in the coalition that contribute the public knowledge good.

Therefore, the only internally stable coalition is the one of size $s = s^{min}$.
This coalition is also the only externally stable because no agent outside the coalition will find it profitable to join it, that is:

$$\pi_c(s^{min}+1) \equiv w-(p-\alpha)g_i+\delta\beta(s^{min}+1)g_i-\theta g_i < w+\delta\beta s^{min}g_i \equiv \pi_{nc}(s^{min})$$

for $\delta < \frac{p-\alpha+\theta}{\beta}$. Considering positive participation cost, θ, this latter condition is always verified in social dilemma scenarios where, by hypothesis, $\delta < \bar{\delta} \equiv \frac{(p-\alpha)}{\beta} < \frac{(p-\alpha)+\theta}{\beta}$.
It is straightforward to demonstrate that the grand coalition $s^* = N$ is always internally and externally stable. In this case and all the agents of the economy will join the coalition during the membership stage of the game. ■

The participation threshold s^{min} determines the minimum number of individuals required in the membership stage to overcome the social dilemma. In this setting the non trivial internally and externally stable coalition is given by the smallest profitable coalition of size s^{min}. If $s < s^{min}$ coalition members do not acquire the good after the membership stage and the $N - s$ agents that do not participate in the coalition maximise their expected payoffs by not investing in the impure good. Similarly, when $s > s^{min}$ at least one member would find it profitable to leave the coalition and internal stability will not be reached.

Comparative statics show that the effects of both the private, α, and the public, β, component of the impure public good on the optimal coalition size are negative.

Interestingly, one can observe an ambiguous effect of a change in α on the aggregate welfare of the coalition equilibrium. The latter can be defined as the sum of the welfare of members and non-members of the coalition $\Pi_{coal} = \Pi_{g,c} + \Pi_{0,nc}$. Notice that:

$$\frac{\partial \Pi_{coal}}{\partial \alpha} > 0 \Rightarrow s^* < N < 2s^* \qquad\qquad \frac{\partial \Pi_{coal}}{\partial \beta} > 0 \ \forall \ N > s^*$$

The aggregate welfare under coalition equilibrium is increasing with the private component α if the size of the population is not too large. The derivative of Π_{coal} with respect to α is given by $\frac{w[3p-2(\alpha-\theta)-\beta\delta N]}{\beta\delta p}$. This derivative is positive if and

only if $s^* < N < 2s^*$. When the grand coalition is not formed, that is $s^* < N$, and total population is not too large, $N < 2s^*$, than the reduction in the number of coalition's members due to an increase in the private component α is dominated by the increased size and aggregate welfare of the agents outside the coalition, ($\Pi_{0,nc}$). When the size of the population is large compared to the coalition size, i.e. $N > 2s^*$, the effect of α on aggregate welfare is negative because the effect of α on the size of the coalition becomes more important. By contrast, the β component always positively impacts the aggregate welfare under coalition equilibrium, via the non-excludable benefits of the public good.

A specific case that deserves mentioning is the case in which the technology of the impure public good is such that the private and public components act as perfect substitutes, for example when $\alpha + \beta = 1$, so that is not possible to increase α, without decreasing β. In the latter case the effect of the increase in the private component α (and the related decrease in public component β) depends on both the cost of participating in the coalition and the price of the impure public good. If the price of the impure public good is sufficiently high compared to the price of the numeraire (normalised to 1), $p > 1 - \theta$, then an increase in the private component α will increase the optimal coalition size. However, when the price of the impure public good is sufficiently low, $p < 1 - \theta$, then the coalition size will decrease.

4 Voluntary contribution to public knowledge goods: theory with social preferences

This section broadens the impure public goods model of coalition formation to the case of social preferences, by that integrating the evidence from social psychology on the role of social preferences in group behaviour. To differentiate the voluntary pools of impure public goods with social preferences from the pools with standard preferences, we designate the former in short by social network coalitions.

As stated in section 2, the empirical literature in social psychology highlights two key dimensions of social preferences that play a prominent role in the involvement in group behaviour: the positive social group identity related to the collective goals realised by the group and social approval for pro-social attitudes. To model the effect

of the positive social group identity related to the belonging of a group that provides public goods, we should consider both the importance of the collective goal related to the aggregate public good provision (cf. for example Wit and Ker, 2002, p. 634) and the fact that group size effects play a role (De Cremer and Leonardelli, 2003). Therefore the social group identity component of the social preference function can be considered as being proportional to the aggregate public good component of the knowledge good produced by the coalition, βG.

For modelling the social approval of the pro-social behaviour in impure public good provision, we suppose that the social approval is proportional to the perceived individual pro-social or altruistic behaviour, as measured by the difference between the individuals contribution to the public knowledge good and his private ancillary benefit from the provision of the public good, that is $(\beta - \alpha)g_i$. This means that when $\beta > \alpha$, the function that gives the social approval of the pro-social behaviour is defined as being positive, while when $\beta < \alpha$, this function is defined as being negative. The latter correspond to a social disapproval of a behaviour that contributes more to the personal private utility (through α) than to the public good (through β).

Social group identity and individual pro-social reputation can be aggregated into the weighted function: $a_i = (1 - \rho)\beta G + \rho(\beta - \alpha)g_i$, with $0 \leq \rho \leq 1$ defined as the weight or subjective value members of a group give to individual pro-social reputation compared to the groups social group identity. If $\rho \to 1$, then agents do not exhibit social preferences for the building of the groups social group identity through the aggregate public good production. Rather they care about signalling and approving/disapproving personal pro-social reputations related to making knowledge publicly available, such as in the case of the community of hackers discussed in section 2. By contrast, when $\rho \to 0$ the agents social preferences are only based on their belonging to a group with a positive social group identity related to the aggregated production of the public good.

Representative agent's preferences are described by the utility function $u_i = \pi_i + v(a_i)$ where function v is assumed to be linear for the sake of simplicity[6]. We

[6]A similar approach can be found in Hollander (1990) when agents have preferences for social approval conceptualised as an emotional activity.

are interested in social dilemma scenarios, which are characterised as above by a situation in which only the pure private commodity (the numeraire) is purchased by agents, but where there exists an incentive for the creation of a coalition where each agent invests its endowment in the impure public knowledge good. Individuals maximise utility $u_i = \pi_i + v(a_i)$ taking into account the behaviour of other agents and the social consequences of their personal actions when acquiring the impure public good. Social dilemma is guaranteed when the marginal material benefit to the non-excludable component is such that $\frac{\bar{\bar{\delta}}}{n} < \delta < \bar{\bar{\delta}}$ with $\bar{\bar{\delta}} = \frac{p-\alpha(1-\rho)-\beta}{\beta}$.

As in the benchmark case, an increase in parameters α and β reduces the marginal per capita cut-off payoff $\bar{\bar{\delta}}$, ρ impacts positively $\bar{\bar{\delta}}$. The assumption of a non-cooperative Nash equilibrium (the social dilemma scenario) implies that individuals outside the coalition will never purchase the impure public good. Assuming such a social dilemma situation with agents that have social preferences, we can derive the optimal size of a coalition able that overcomes the non-cooperative equilibrium by comparing the utility of a member of the coalition acquiring the impure public good with the utility of a member that prefers not to acquire this type of good. Proceeding as in the previous section we obtain:

$$\Delta u_c \equiv u_{g,c} - u_{0,c} = g_i\{s\beta[(1-\rho)+\delta] + \alpha - p + \rho(\beta - \alpha)\} \tag{7}$$

Proposition 2 *When agents exhibit social preferences for social identity and pro-social attitudes, $a_i = (1-\rho)\beta G + \rho(\beta - \alpha)g_i$, the optimal stable coalition is given by:*

$$s^{**} = \frac{p - \alpha + \theta + \rho(\alpha - \beta)}{\beta[(1-\rho)+\delta]}$$

The coalition of size s^{**} is the sub-game perfect Nash equilibrium of the voluntary coalition game with social preferences. When $s^{**} \geq n$ then a grand coalition with $s^{**} = n$ forms.

Proof. *To demonstrate that a coalition of size s^{**} is the sub-game prefect Nash equilibrium of the game with social preferences we proceed as in proposition 1. Solving the term in brackets in (7) we find the minimum profitable coalition size $s^{min} = \frac{p-\alpha+\theta+\rho(\alpha-\beta)}{\beta[(1-\rho)+\delta]}$. Assume a coalition with a number of members equal to the*

minimum profitable size. If one member defects then no member will find it convenient to purchase the impure public good and the public good will be not produced by the social network. Comparing the utility of an agent when at least one member leaves the coalition with the utility of a member when the minimum profitable size is reached allows to observe that $u_c(s^{**}) \geq u_{nc}(s^{**}-1)$*. This condition is verified if and only if* $s \geq s^{**} = \frac{p-\alpha+\theta+\rho(\alpha-\beta)}{\beta[(1-\rho)+\delta]}$*. However, if the coalition size is strictly greater than* s^{**} *then at least one member will find it profitable to leave the coalition. Using* s^{**} *we easily observe that* $u_c(s^{**}) = u_{nc}(s^{**}-1)$*. It follows that the only internally stable coalition is the one with size* s^{**}*.*

External stability requires that $u_c(s^{**}+1) < u_{nc}(s^{**})$*, meaning that an agent outside the coalition will never find it profitable to join the coalition during the membership stage. Comparing* $u_c(s^{**}+1) \equiv w-(p-\alpha)g_c+\delta\beta(s^{**}+1)g_c-\theta g_c+(1-\rho)\beta(s^{**}+1)g_c+\rho(\beta-\alpha)g_c < w+\delta\beta s^{**}g_c+(1-\rho)\beta(s^{**})g_c \equiv u_{nc}(s^{**})$ *we observe that the coalition* s^{**} *is externally stable if and only if* $\delta < \frac{p-\alpha(1-\rho)-\beta+\theta}{\beta}$*. This condition is always satisfied under social dilemma with social preferences,* $\delta < \bar{\bar{\delta}}$*.*

As before, the grand coalition is the trivial outcome of the game when $s^{**} \geq N$*, that is, the entire population will join the coalition during the membership stage of the game.* ■

The optimal coalition of size s^{**} coincides with the smallest profitable coalition also for agents with social preferences. However, as the discussion in section 2 of the stylised facts of coalition formation in FLOSS communities shows, the social preferences produce contrasted effects on the facility of coalition formation. Using the model of coalition formation with social preferences, it is possible to show that assuming social preferences indeed not necessarily implies that the stable coalition size will be smaller than in the case of standard preferences.

Corollary 1 *If* $\beta < \alpha$*, depending on weight* ρ*, the optimal coalition size created by agents with social preferences can be larger or smaller with respect to the optimal size of a coalition formed by individuals with standard preferences:*

- *if* $\rho > \bar{\rho} \Rightarrow s^{**} > s^{*}$
- *if* $\rho < \bar{\rho} \Rightarrow s^{**} < s^{*}$

with $\bar{\rho} = \frac{p-\alpha+\theta}{p-\alpha+\delta(\alpha-\beta)+\theta}$. *If* $\beta \geq \alpha$ *the coalition size under social preferences is always smaller than (or equal to) the coalition size under standard preferences,* $s^{**} \leq s^{*}$.

Comparing the optimal coalition size under social and standard preferences, we observe that $s^{**} \equiv \frac{p-\alpha+\theta+\rho(\alpha-\beta)}{\beta[(1-\rho)+\delta]}$ can be larger of $s^{*} \equiv \frac{p-\alpha+\theta}{\beta\delta}$ if and only if $\rho > \bar{\rho} \equiv \frac{p-\alpha+\theta}{p-\alpha(1-\delta)-\beta\delta+\theta}$. If the return to the public component is sufficiently small, e.g. $\beta < \alpha$, and the weight to pro-social reputation is such that $\rho > \bar{\rho}$, then the coalition formation will be more difficult to form for agents with social preferences compared to the scenario in which agents exhibit standard preferences. By contrast, if the weight for social group identity related to the collective action is sufficiently high, so that $\rho < \bar{\rho}$, then the optimal coalition size for agents with social preferences will be smaller with respect to the coalition composed by agents with standard preferences, $s^{**} < s^{*}$.

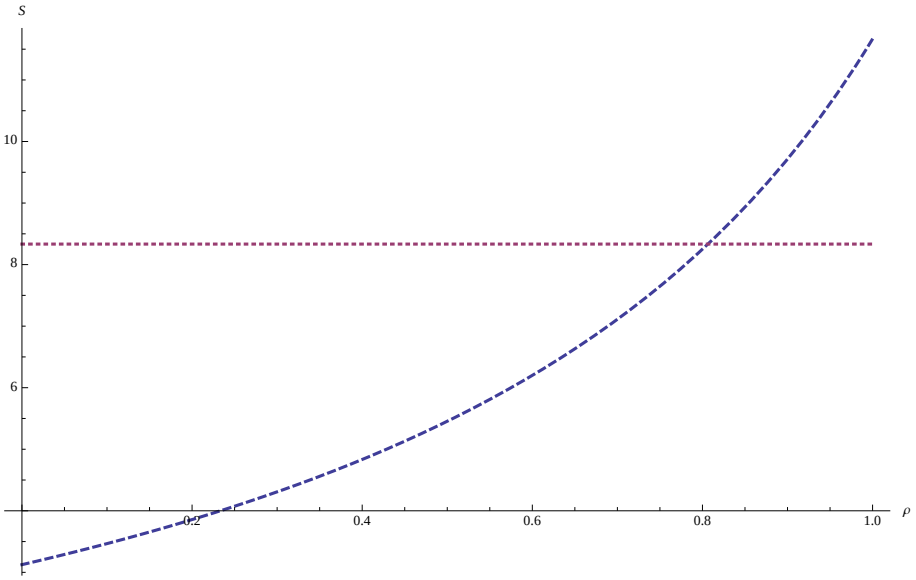
As stated in the corollary, if the technology of the impure knowledge good is such that $\beta < \alpha$, then the effects of the social preferences are ambiguous: depending on the weight for group identity or comparative approval, it might or might not make coalition formation easier. This situation is represented in figure 2, which graphically describes the optimal coalition size for different values of the comparative weight given to social approval for pro-social attitudes[7] ρ, compared to the weight given to the social group identity $(1-\rho)$. Figure 2 illustrates that if the weight for pro-social reputation is sufficiently high, $\rho > 0.785$, then the optimal coalition size for agents with social preferences (dashed curve) will be larger than the optimal size for agents with standard preferences (dotted line), and therefore the coalition will be more difficult to form. The larger the weight for the individual pro-social reputation in the social preferences function compared to the social group identity (e.g. hackers communities), the higher the optimal coalition size under social preferences will be.

As stated above, if the technology of the networked good is such that $\beta < \alpha$, depending on the weight for group identity or comparative approval, then the effects of the social preferences is ambiguous: it might or might not make coalition formation easier.

The effect of α on the aggregate welfare when a coalition equilibrium is reached

[7] We use the following parameters' values for this numerical example: $\beta = 0.2, \alpha = 0.6, p = 1.5; \theta = 0.1; \delta = 0.6$.

Figure 2: Optimal size under standard and social preferences



is the same as the one observed in the case of standard preferences. As in the benchmark case, the total welfare is increasing in the private component α if the size of the population is not too large, that is $s^{**} < N < 2s^{**}$. Otherwise it is negative.

5 Application: Heterogeneous Social Preferences

In this section we assume a heterogeneous population in which individuals differ in terms of their social preferences for social group identity and social approval of individual pro-social attitudes. More precisely, we assume two types of agents who differs only in terms of their social preference function: agents exhibit either more weight for individual pro-social reputation $\bar{\rho}$ or they exhibit more weight for social group identity related to the provision of the public good by the coalition $\underline{\rho}$, with $\bar{\rho} > \underline{\rho}$. To illustrate this situation based on the categories of the FLOSS surveys (cf. table 1), one can think for instance of a population of developers in which some members participating in social networks have strong preferences for their private pro-social reputation, such as in the case of a community of hackers developing open source software, while other agents exhibit social preferences for building the groups social identity through their involvement in the public good provision, such

as social programmers or professionals. The questions that we want to investigate in this context are the following: what is the optimal coalition size when the social networks are composed of agents with different social preference functions? Is the coalition formed by heterogeneous agents stable? Does this type of coalition form when homogeneous groups are not able to implement a community composed by agents with the same preferences?

Assume that individuals are rational. They anticipate the strategy of other agents when maximizing their utility function in investing time or monetary endowment in the impure public good. The choice of participation is made simultaneously by both type of agents. An important difference with the case of homogeneous preferences is that agents know that if they do not participate in the coalition they will still be able to free ride on the public good produced by a coalition formed by rational agents of the other type. Therefore, the payoff when not contributing to the coalition is higher, compared to the case in which agents exhibit the same social preferences and do not contribute to the coalition.

In order to derive the minimum profitable mixed coalition size when the social preferences are heterogeneous, we have to solve the following system of equations:

$$\begin{cases} \Delta\bar{u} = w_i - [(p-\alpha) - \theta + \beta\delta s^m]g_i + (1-\bar{\rho})\beta s^m g_i + \bar{\rho}(\beta-\alpha)g_i - (w_i + \beta\delta\underline{s}g_i) \\ \Delta\underline{u} = w_i - [(p-\alpha) - \theta + \beta\delta s^m]g_i + (1-\underline{\rho})\beta s^m g_i + \underline{\rho}(\beta-\alpha)g_i - (w_i + \beta\delta\bar{s}g_i) \end{cases} \tag{8}$$

with $\bar{s}$ and $\underline{s}$ defining the number of members respectively with social preferences $\bar{\rho}$ and $\underline{\rho}$ participating in the heterogeneous coalition. The sum $s^m = \bar{s} + \underline{s}$ defines the size of the coalition when agents exhibit heterogeneous social preferences and the choice of participation is made simultaneously. When maximising the net utilities $\Delta\bar{u}$ and $\Delta\underline{u}$, rational agents know the number of members of the other groups willing to participate in the mixed coalition. Therefore, they also know their personal material benefit of the free riding when not participating.

Solving the system defined in (8) allows to derive the minimum profitable mixed coalition size formed by agents with different social preferences as well as the share

of each group within the coalition. We find that:

$$\bar{s} = \frac{\beta(\underline{\rho} - \bar{\rho}) - \alpha\delta(1 - \bar{\rho}) + (p + \theta)(\delta + \bar{\rho} - \underline{\rho}) - \beta\delta\bar{\rho}}{\beta\delta(2 + \delta - \bar{\rho} - \underline{\rho})} \tag{9}$$

$$\underline{s} = \frac{\beta(\bar{\rho} - \underline{\rho}) - \alpha\delta(1 - \underline{\rho}) + (p + \theta)(\delta + \underline{\rho} - \bar{\rho}) - \beta\delta\underline{\rho}}{\beta\delta(2 + \delta - \bar{\rho} - \underline{\rho})} \tag{10}$$

$$s^m = \frac{2(p - \alpha + \theta) + (\alpha - \beta)(\underline{\rho} + \bar{\rho})}{\beta\delta(2 + \delta - \bar{\rho} - \underline{\rho})} \tag{11}$$

Proposition 3 *(i) Assume $\alpha > \beta$. If $\rho^- < \bar{\rho} \leq \rho^+$ then the mixed coalition of size s^m, with $\bar{s}$ members of type $\bar{\rho}$ and $\underline{s}$ members of type $\underline{\rho}$, with $\underline{s} < \bar{s}$, is the sub-game Nash equilibrium of the game. (ii) Assume $\beta > \alpha$. If $\underline{\rho} < \frac{p-\alpha+\theta}{\beta-\alpha}$ and $\rho^- < \bar{\rho} \leq \rho^+$ then same result as (i); if $\underline{\rho} \geq \frac{p-\alpha+\theta}{\beta-\alpha}$ then the mixed coalition of size s^m is not stable.*

Proof. *We define $\rho^- = \frac{\alpha\delta+(p+\theta)(\underline{\rho}-\delta)-\beta\underline{\rho}}{p+\delta(\alpha-\beta)+\theta-\beta}$ and $\rho^+ = \frac{(\beta-\alpha)\delta(1-\underline{\rho})}{p+\theta-\beta} + \delta + \underline{\rho}$. Notice that if $\alpha > \beta$ then $\rho^- < \rho^+$. First, assume that $\alpha > \beta$. In order to study the stability of the minimum profitable mixed coalition s^m we have to prove that simultaneously both type of members have no incentive to leave the coalition and that agents outside the coalition have no incentive to join it. Comparing the utility of an agent with social preferences $\underline{\rho}$ when one member leaves the coalition with the utility of a member with social preferences $\underline{\rho}$ when the size s^m is reached, allows to observe that $\underline{u}_c(s^m) - \underline{u}_{nc}(s^m - 1) = \frac{g[\beta(\underline{\rho}-\bar{\rho})-\alpha\delta(1-\bar{\rho})+(p+\theta)(\delta+\bar{\rho}-\underline{\rho})-\beta\delta\bar{\rho}]}{\beta\delta(2+\delta-\bar{\rho}-\underline{\rho})} \equiv g\bar{s} \geq 0$ if and only if $\bar{\rho} \geq \rho^-$. If the coalition size is strictly greater than s^m, then at least one member will find it profitable leave the coalition. Using (11) we observe that $\underline{u}_c(s^m) = \underline{u}_{nc}(s^m - 1)$ when $\bar{\rho} = \rho^-$. Knowing that $\bar{\rho} > \underline{\rho}$ by assumption, it follows that only when agents exhibit social preferences such that $\bar{\rho} > \rho^-$ the mixed coalition s^m is profitable and internally stable for agents with social preferences $\underline{\rho}$. This coalition is also externally stable for agents with social preferences $\underline{\rho}$. It is easy to prove that the external stability condition is verified, that is, $\underline{u}_c(s^m + 1) < \underline{u}_{nc}(s^m)$ for all $0 \leq \underline{\rho} < \bar{\rho}$.*

Similarly, for an agent with social preferences $\bar{\rho}$, internal stability is reached if and only if $\bar{u}_c(s^m) - \bar{u}_{nc}(s^m - 1) = \frac{g[\beta(\bar{\rho}-\underline{\rho})-\alpha\delta(1-\underline{\rho})+(p+\theta)(\delta+\underline{\rho}-\bar{\rho})-\beta\delta\underline{\rho}]}{\beta\delta(2+\delta-\bar{\rho}-\underline{\rho})} \equiv g\underline{s} \geq 0$. This condition is verified when social preferences are such that $\bar{\rho} \leq \rho^+$. When this latter condition is met the mixed coalition s^m is profitable and internally stable for agents

with social preferences $\bar{\rho}$. Also for this type of agents the coalition is externally stable, $\bar{u}_c(s^m + 1) < \bar{u}_{nc}(s^m)$ for all $\underline{\rho} < \bar{\rho} \leq 1$. Therefore, the mixed coalition s^m is internally and externally stable coalition and it is the sub-game Nash equilibrium of the game if and only if social preferences are such that $\rho^- < \bar{\rho} \leq \rho^+$.

Assume now that $\beta > \alpha$. In this case it is possible that the mixed coalition is not stable also when $\rho^- < \bar{\rho} \leq \rho^+$. If the preferences of the agent $\underline{\rho}$ are such that $\underline{\rho} \geq \frac{p-\alpha+\theta}{\beta-\alpha}$, than $\rho^- \geq \rho^+$ and internal stability is never reached. By contrast, if $\underline{\rho} < \frac{p-\alpha+\theta}{\beta-\alpha}$ and $\rho^- < \bar{\rho} \leq \rho^+$ the mixed coalition is internally and externally stable, for the reasons discussed above.

Notice that in the proof we are assuming that no agent is able to form a coalition with individuals of the same social preferences type, that is $\underline{s} < s^{**}$ for $\underline{\rho}$ and that $\bar{s} < s^{**}$ for $\bar{\rho}$. This means that when the coalition is not formed the public knowledge good is not produced. Supposing that agents with the same social preferences are able to form coalitions with agents of the same type does not change the results: mixed coalitions are internally and externally stable also when agents can free ride over the contribution of a coalition formed by individuals with different social preferences. ■

As in the model with homogeneous population an increase (decrease) in the private return of the impure public good α leads to smaller (bigger) optimal size s^m of the mixed coalition. In other words, an increase (decrease) in private return generated by the impure good makes coalition formation easier (more) difficult. For example, in the FLOSS case discussed in section 2, for a heterogeneous coalition composed of social programmers (with a comparative high social preference for social group identity, that is a low $\underline{\rho}$) and hackers (with a comparative high social preference for the approval of individual pro-social behaviour related to freeing the knowledge, that is a high $\bar{\rho}$), this means that a decrease in private return will lead to bigger coalition sizes. This can be explained by two contrasting effects: the decrease in private return leads to a lower willingness to participate in small coalitions by the social programmers, but the same decrease leads to a higher willingness to participate in small coalitions by the hackers. However, the social group identity effect dominates the impact on the individual reputation in the heterogeneous coalition.

Interestingly enough, it is possible to observe that when a social network with

members characterised by the same type of preferences is not formed because the minimum profitable coalition size is not reached, a mixed coalition with both type of members might be implemented. Imagine for example that in the membership stage agents of type j, with $j \in \{\underline{\rho}; \bar{\rho}\}$ and s^{**} the optimum coalition size of the homogeneous coalition, do not form any coalition because $s^j < s^{**}$ when $s^j \in \{\underline{s}; \bar{s}\}$. However, it is possible that a mixed coalition $s^m = \bar{s} + \underline{s}$ is implemented and is stable even when $s^j = \bar{s}$ for $j = \bar{\rho}$ and $s^j = \underline{s}$ for $j = \underline{\rho}$. This result is driven by the fact that agents might find it profitable to collaborate with other agents with different social preferences in order to reach the minimum profitable coalition size that they are not able to reach when creating a community of individuals with the same social preferences. As can be seen from proposition 3 (ii), this outcome, to be valid in general, requires the presence of agents with strong social preferences for social group identity related to the public good production by the coalition (that is low ρ).

Finally, when the private return of the impure public good is relatively low compared to its public benefit, $\alpha < \beta$, it is possible that mixed social networks fail to form (second part of proposition 3 (ii)). In particular, if no agent has strong preferences for social group identity or, in other words, if the two groups of agents have very similar and high preferences for individual pro-social reputation, then the mixed community might be not stable. Assume for instance that $\underline{\rho}$ is sufficiently high: $\underline{\rho} > \frac{p-\alpha+\theta}{\beta-\alpha}$. This implies that both social groups have strong social preferences for pro-social attitudes, that is they have strong disapproval of self-regarding attitudes (imagine for instance two groups of hackers with similar preferences in a context in which the private return to the impure good is low). If these conditions are met the mixed coalition is not an equilibrium of the game. In this case the agents prefer to form a coalition only with individuals of the same type.

These results from the theory of coalition formation with a heterogeneous population are supported by the general trends of the FLOSS-US 2003 survey as analysed by David and Shapiro (2008).

Table 2 gives the matrix for developers' movements among projects of different membership size. In reading these data, it is reasonable to make the hypothesis that on average developers in their first project derive higher personal learning benefits

Table 2: Movement of developers from first projects to current/most recent projects

Current/most recent project	First project		
	Small 1-2	Medium 3-29	Large > 29
Large	51	42	32
Medium	30	29	22
Small	13	7	18
Pearson chi-squared (4)	7.96		
Prob>chi-squared	0.09		

Source: David and Shapiro (2008), p. 389, Table 9, Panel B (Data from US-FLOSS 2003 survey, for the sub-set of 244 respondents whose current and first project were not the same)

(higher α), compared to the involvement of these same persons in their second and/or most recent project, especially if they were involved in their first project in projects of small and medium size. Indeed as stated by David and Shapiro (2008, p. 390) "while these communities may well be regarded as sites of software skills development and organisational know-how that foster open-source production at large, it appears that they are also absorbing, and enhancing the programming abilities of developers who made a start on individual projects, undertaken without formal instruction in their capacity as students". As can be seen in table 2, the larger projects function as attractors in the overall circulation dynamics of those who change projects. Therefore, on average, agents in second stage projects (with lower private return on the impure public good) show a higher probability to go to larger FLOSS projects, as predicted by the theoretical model. In particular, this result is valid both for agents that were involved in the first stage in small groups, which are likely to be homogeneous, or in medium groups, which include both homogeneous and heterogeneous cases. The first case corresponds to homogeneous groups that extend to heterogeneous groups to reach the stable coalition size. The second case corresponds to the increase in optimum group size for heterogeneous coalitions, when alpha is decreasing. For the agents that were involved in the first stage in large groups, it is more difficult to draw conclusions as the private learning effects are likely to be less pronounced or at least unequally distributed over the population.

Based on the analysis of these trends, it would be worthwhile, to further test these results by combining the circulation data with the data on the agents belonging to one of the various clusters presented in table 1. This will be the topic of further

research, along with the collection of new primary data through large-scale semi-structured interviews of data sharing and materials sharing in the field of knowledge generation on and use of genetic resources.

6 Conclusion

This paper has analysed the ambiguity of private ancillary benefits in fostering coalition formation with publicly available knowledge goods. Such ancillary benefits are considered as an important driver of the proliferation of pooled knowledge assets in social networks. Ancillary benefits that have been widely studied in the literature are of two kinds: (1) direct private benefits such as learning of new competence or higher citation rates for researchers; (2) satisfaction of social preferences such as group belonging, group identity, pro-social individual reputation and status.

In the current literature on coalition formation for public good provision, the effect of these two kinds of ancillary benefits is mostly considered to be positive on the facility to form the coalition. No explanation of the ambiguous effects is provided (that is an explanation that in some cases the ancillary benefits make the coalition formation easier and in some case more difficult). To build a more general theoretical model, this paper integrated the theory of impure public goods and a social psychological model of group related social preferences into coalition theory. This allowed us to show the contrasted effects of social group identity and social approval/disapproval of individual pro-social attitudes on the coalition formation. The presence of agents giving high value to their individual pro-social reputation within a social network can make coalition formation in some situations more difficult and less likely.

Acknowledgements

We gratefully acknowledge co-funding from the European Commission, under the contract of three PF7-projects, including LowInputBreeds (Grant agreement 222623), BIOMOT (grant agreement 282625) and GENCOMMONS (ERC grant agreement 284) and co-funding from the National Science Foundation (MIS Global Science Commons).

References

[1] Andreoni, J. 1988. Privately provided public goods in a large economy: the limits of altruism. Journal of Public Economics 35(1), 57-73.

[2] Barret, S., 1994. Self-enforcing international environmental agreements. Oxford Economic Papers 46: 878-894.

[3] Benabou, R. and J. Tirole. 2003. Intrinsic and extrinsic motivation. Review of Economic Studies 70(3): 489-520.

[4] Benkler, Y. 2006. The Wealth of Networks. Yale UP.

[5] Bergstrom, T. et al. 1986. On the private provision of public goods. Journal of Public Economics 29: 25-49.

[6] Blader, S. L and T. R. Tyler. 2009. Testing and extending the group engagement model: linkages between social identity, procedural justice, economic outcomes, and extrarole behaviour. Journal of Applied Psychology 94(2): 445464.

[7] Bolton, G.E., Katok, E. and Ockenfels, A. 2005. Cooperation among strangers with limited information about reputation. Journal of Public Economics 89(8): 1457-1468.

[8] Carraro, C. and D. Siniscalco. 1993. Strategies for the international protection of the environment. Journal of Public Economics 52(3): 309-328.

[9] Cornes, R. and T. Sandler. 1984. Easy Riders, Joint Production, and Public Goods. The Economic Journal 94(3): 580-598.

[10] Cornes, R. and T. Sandler. 1994. The Comparative Statics Properties of the Impure Public Good Model. Journal of Public Economics 54(3): 403-421.

[11] David, P. A. 2005. Koyaanisqatsi in Cyberspace: The Economics of an Out-of-Balance Regime of Private Property Rights in Data and Information. In International Public Goods and Transfer of Technology under a Globalized Intellectual Property Regime, ed. K. E. Maskus and J. H. Reichman, 79-120. Cambridge University Press.

[12] David, P. A. 2008. The historical origins of 'Open Science': An essay on patronage, reputation and common agency contracting in the scientific revolution. Capitalism and Society 3(2): 5.

[13] David, P. A. and J.S. Shapiro. 2008. Community-based production of open-source software: what do we know about the developers who participate ? Information Economics and Policy 20(4): 364-398.

[14] De Cremer, D. and G. J. Leonardelli. 2003. Cooperation in social dilemmas and the need to belong: the moderating effect of group size. Group Dynamics: Theory, Research, and Practice 7(2): 168174.

[15] Giddens, A. 1999. Runaway World: How Globalization is Reshaping Our Lives. London: Profile.

[16] Hess, C. and E. Ostrom. 2007. Understanding Knowledge as a Commons. MIT Press.

[17] Joireman, J. et al. 2006. Good citizens to the end? It depends: empathy and concern with future consequences moderate the impact of a short-term time horizon on organizational citizenship behaviors. J Appl Psychol. 91(6): 1307-1320.

[18] Kosfeld, M. et al. 2009. Institution Formation in Public Goods Games. American Economic Review 99(4): 1135-1355.

[19] Kotchen, M. J. 2006. Green Markets and Private Provision of Public Goods. Journal of Political Economy 114(4): 816-834.

[20] Kotchen, M. J. 2007. Equilibrium Existence and Uniqueness in Impure Public Good Models. Economics Letters 97: 9196.

[21] Kotchen, M. J. 2009. Voluntary Provision of Public Goods for Bads: A Theory of Environmental Osets. The Economic Journal 119: 883-899.

[22] Lepine, J. A et al. 2002. The nature and dimensionality of organizational citizenship behavior: a critical review and meta-analysis. J Appl Psychol 87(1):52-65.

[23] Lessig, L., 2001. The Future of Ideas. Random House.

[24] Lessig, L. 2008. Remix: Making Art and Commerce Thrive in the Hybrid Economy. Penguin Press.

[25] Penner, L. A. and M. A. Finkelstein. 1998. Dispositional and structural determinants of volunteerism. Journal of Personality and Social Psychology 74(2): 525-537.

[26] Raymond, E. S. 1999. The Cathedral & the Bazaar. O'Reilly Media.

[27] Reichman, J. H. and P. F. Uhlir. 2003. A contractually reconstructed research commons for scientic data in a highly protectionist intellectual property environment. Law and Contemporary Problems 66: 315.

[28] Sandler, T. and D. G. Arce. 2007. New face of development assistance: public goods and changing ethics. Journal of International Development 19(4):527-544.

[29] Suurmond, G., Swank, O.H. and Visser, B. 2004. On the bad reputation of reputational concerns. Journal of Public Economics, 88(12), 2817-2838.

[30] Vicary, S. 1997. Joint production and the private provision of public goods. Journal of Public Economics 63(3), 429-445.

[31] Vicary, S., 2000. Donations to a public good in large economy. European Economic Review 44(3): 609-618.

[32] Wit, A. P and Ker, N. L. 2002. Me versus just us versus us all categorization and cooperation in nested social dilemmas. Journal of Personality and Social Psychology 83(3): 616-637