

The Effect of Privacy Concerns on Social Network Formation

Alexia Gaudeul¹ and Caterina Giannetti²

July 4, 2017

Abstract

We study the impact of revealing personal information on the selection of partners when forming individual networks. In our experiment, a “contribution phase” is followed by a “network phase”: individuals first decide whether to contribute and then decide with whom to form a link. An exchange of contributions between individuals occurs only if a bilateral link is established. We identify the effect of revealing one’s name on the probability of link formation by letting individuals choose to signal – at the linking stage – their willingness to later reveal their name. Our results indicate that this decision significantly affects partner selection and the consequent structure of the network. Individuals who reveal their name build up smaller individual networks but attain higher profits. Incurring privacy costs by revealing personal information is compensated with a higher probability to establish mutual links with contributors.

Keywords: privacy, social networks, public goods, trust

JEL Classification: D12, D85

¹*Corresponding author.* Department of Economics, Georg-August-Universität, Göttingen, Germany. Email: alexia.gaudeul@wiwi.uni-goettingen.de

²Department of Economics and Management, University of Pisa, Italy.

The authors thank the Associate Editor, three anonymous referees, Pietro Battiston, Francesco Feri, Alisa Frik, Michael Kammer, Oliver Kirchkamp, Alessio Moneta and participants at the ZEW Conference on Information and Communication Technologies in Mannheim for useful comments and suggestions. The experiment was funded by the Max Planck Institute of Economics in Jena. The authors also thank Ria Stangneth for her excellent work in programming the experiment and Clara Dubois for her help in running sessions in the lab.

Having given both to the just and the unjust power to do what they will, let us watch and see whither desire will lead them. (The Ring of Gyges, Plato)

1 Introduction

Individuals disclose a huge amount of information about themselves on social networks about their hobbies, tastes in music, or favorite books. They also disclose personally identifiable information such as their name, birthday and location. Surprisingly, they fill in and update this information by themselves over the years. Apparently, the perceived benefits of these disclosures – which allow other participants of social networks as well as firms to better know their preferences – outweigh the costs that could arise if this information is misused – the privacy risk (for a review, see [Acquisti et al. \[2016\]](#)).

In this paper we show that disclosing private information on social networks serves an instrumental purpose, as people who reveal more personally identifiable private information are better able to connect with others and develop fruitful relations. We draw inspiration from the current debate about the conflict between protecting privacy online and protecting Internet users from exploitation and anonymous attacks. Many social networks allow their members to maintain anonymity. However, although this (online) anonymity gives individuals the freedom to disclose more about themselves, it can also lead them to “toxic disinhibition” (e.g. posting inflammatory messages in public forums, see [Suler \[2004\]](#), [Belk \[2013\]](#)). Various systems can be used to regulate communication and maintain communities, from establishing reputation systems to excluding misbehaving participants, as well as establishing other barriers to be overcome in order to gain access to communities, such as requiring an invitation by current members, or payment of a fee, or inflicting “hazing” on newcomers [[Friedman and Resnick, 2001](#), [Bacharach and Board, 2002](#)]. However, these systems rely on being able to associate a person with their action. For this reason, content providers, social networking platforms and community sites are moving towards lifting the veil of anonymity online. Against this trend, however, governments worldwide are considering new privacy regulations that would restrict the processing and disclosing of personal data (see e.g. the European Privacy Directive). For example, Facebook’s “real name policy”, which requires its users to use their real names, has been the subject of legal controversy both in the US and in Germany. Despite this debate on the desirability of being able to identify people online, to our knowledge no systematic studies have been carried out to evaluate how revealing personally identifiable information affects the process of network formation. The aim of our study is to fill this gap; ours is the first economic experiment that considers how privacy concerns affect network formation. We intend to show how being able (but not forced) to reveal one’s identity allows people to develop relationships in a context that is otherwise anonymous and disincarnate. We investigate whether such an endogenous name revelation mechanism can work on its own, without having to rely on

further community management devices such as punishment, exclusion or tools allowing reputation building.

There are significant advantages in using experimental research methods instead of doing observational research, and experiments are proving to be an invaluable tool for studying network formation [Choi et al., 2016]. This is because it is difficult (if not impossible) to infer causal relationships from field data as there are many confounding features in a network environment, such as simultaneous influences or measurement errors. However, when it comes to studying privacy issues in the laboratory, the main disadvantage of experimental methodology is that it is difficult to generate private information that can be safely used in the laboratory and is also sufficiently sensitive for the participants. For example, it is not feasible to reveal credit card numbers as they could be used outside of the laboratory for malign purposes.

We therefore generated sensitive private information by structuring our experiments into two phases. The “contribution phase” was used to generate private and sensitive information, while the “network formation phase” came later. In the contribution phase, individuals had to decide whether to contribute or not. Contributing meant that anyone with whom the individual later established a mutual link would receive a multiple of the contribution, while not contributing meant that he/she would receive nothing. This first phase allowed us to generate a sensitive type of information in a synthetic way: whether the individual contributed or not, which could later be disclosed in the laboratory. Moreover, this type of decision—contributing or not—resembles the type of decisions that are made in social networks. The more people are active in a network—which requires some effort—the more valuable a relationship with them may be. However, some individuals may free ride on the activity of others without revealing anything of interest about themselves and their activity (see Goyal et al., 2014; Graham, 2014). Since it is not always possible to see how much another person reveals about him/herself or how beneficial a link with him/her will be before adding them to one’s list of “friends” on Facebook, linking with someone may involve careful thought about whether it is worth revealing one’s personal information and network of friends to him or her.

In the network formation phase, which occurred after the contribution phase, the participants in our experiment were assigned to small groups and were asked to decide with whom to create a link. A multiple of their contribution in the first phase was exchanged between two individuals in the group only if both expressed the desire to form a link with each other (i.e. only for reciprocal links, Jackson and Wolinsky, 1996). For simplicity, and unlike on Facebook for example, link solicitations were made by everyone at the same time without knowing others’ decisions. From the point of view of the participants, the purpose of the game was therefore to form mutual links with as many contributing members as possible, and to avoid fruitless links with non-contributing members.

We identify the impact of privacy concerns on the linking choice of participants by varying across treatments the type of information individuals could reveal about themselves during the network formation phase, while keeping the matching technology constant. We call “signal” the informa-

tion the individuals could transmit to other participants. The signal simply indicated a decision made by an individual (depending on the treatment, to reveal their name, to reveal a fake name, to pay a fee), but could affect the way this individual might be perceived by others. Another way to call this signal is a “tag”, i.e. a signal that connotes an attribute of an individual or of a groups (see [Bravo et al., 2012](#) and [Axelrod et al., 2004](#)).

In the *Baseline* treatment, individuals could not signal, and thus chose their links without knowing anything about their counterparts. In the *Names* and in the *Names+Info* treatments, participants could decide – before choosing their links – whether to later disclose their real name. A signal placed next to their ID number at the linking stage indicated their decision to do so. Crucially, we avoid possible confounding factors linked with differences in names (e.g. as a signal of gender, social class or nationality), as the signal only indicated the decision to reveal one’s name, while the name of the person was disclosed only after the linking decisions had been made.¹ Participants in the *Names* treatment only learned the identity (if revealed) of those they linked with at the end of the experiment — thus making individual identification of contributions impossible in most cases — while participants in the *Names+Info* treatment additionally learned both the identity (if revealed) and the *individual* contributions of each member of their group. Therefore, participants could incur a non-monetary cost (i.e. a privacy cost due to the revelation of personal information) in the *Names* and *Names+Info* treatments if they decided to identify themselves (directly in *Names+Info*, possibly by inference in *Names*). This privacy cost was potentially higher for non-contributors,² and varied across participants and across treatments, depending on their sensitivity to the information produced in the laboratory. Some individuals may really dislike (or not care at all) about their name being associated to non-contribution, and they may be more or less sensitive to the certainty that their contribution will be known to others.

We ran two additional treatments to evaluate the possible impact of two ways in which signals could affect link formation and potentially help develop reciprocal links without generating a privacy cost. In the *Fake* treatment, individuals could choose a nickname and signal their decision to do so at the linking stage. In the *Fee* treatment, individuals could signal their decision to pay a fee before the linking stage. In both of these treatments, individuals thus observed a signal that could vary across group members (which is not the case in the *Baseline* treatment), but this signal did not have any privacy consequences (unlike in the *Names* and *Names+Info* treatments).

In summary, the contribution of this paper is to provide a simple experimental analysis of the effect of privacy concerns as an endogenous force in the formation of individual networks. The purpose of this experiment is to identify whether individuals may trade personal private information in order to obtain more productive social networks. In particular, we are interested in determining whether the decision to reveal one’s name positively affects the probability of forming productive

¹As a shortcut, we will say that an individual *signals* if he decides to reveal his name, and *does not signal* if he decides not to reveal his name. Of course, in a wider sense of the term, not signaling is also a signal.

²The name and contribution constitute personal information as defined by the EU Data Protection Directive 95/46/EC, where it is stated that personal data is any information relating to an identified or identifiable natural person.

links with other members of the network. In order to clearly identify those effects, we kept the structure of the network as simple as possible— only direct mutual links mattered—and we compare treatments in which our subjects could signal their decision to reveal their real name with treatments which involve other types of signals that do not imply any privacy cost.

Our results suggest that participants regarded their contribution as sensitive private information (the share of those who revealed their name but did not contribute was lower in treatments *Names* and *Names+Info*), and cared about obtaining social approval (the share of those who revealed their name and contributed was higher in treatment *Names+Info*). Importantly, the decision to reveal one's name significantly and positively affected the probability of establishing a link and with whom, as those who contributed and revealed their name were more circumspect about whom they submitted link proposals to. Their networks were thus substantially smaller, but they were able to obtain higher profits than contributors who did not reveal their name by carefully selecting their links according to the signals of the other participants. However, while contributors were more likely to reveal their name and more likely to only link with others who revealed their name, a number of non-contributors also revealed their real names and thus attracted links from contributors. Therefore, contributors were not fully able to protect themselves from exploitation by the non-contributors who revealed their names. We conclude that simple privacy mechanisms (such as being able to give one's real name) significantly affect the selection of partners and improve the profitability of maintaining a network of relations.

The structure of the paper is as follows: Section 2 reviews the existing literature, Section 3 describes the design of the experiment along with the matching technology and the payoff structure, Section 4 describes the data and presents the results of the univariate analysis, Section 5 describes the econometric analysis and presents results from dyadic regressions. Finally, Section 6 summarizes and concludes our argument.

2 Literature review

Much of the research about privacy has focused on how firms gain personal information from individuals by offering them discounts, better targeted ads and personalized services in exchange (see, for example, [Acquisti and Grossklags, 2005](#), [Grossklags and Acquisti, 2007](#); [Goldfarb and Tucker, 2011](#)). The main objective of those studies has been to investigate the economic consequences of consumers' decisions to protect or share their personal information. From this strand of literature, it emerges that social media networks have facilitated a culture of disclosure. Privacy choices are in the hands of consumers, but many of them lack the awareness and technical ability to protect their personal information. However, none of those papers has considered the crucial role of disclosing personal and private information as a driving force in network formation. We contribute to the privacy literature by showing that the disclosure of personal and private information also serves an instrumental purpose; it helps to build social networks as people who reveal personally identifiable private information in social networks are better able to connect with others and

develop fruitful relations. Therefore, even those who dislike revealing personal information and know how to protect it may want to disclose it.

Previous economic experiments on privacy consider the impact of identification in standard games such as the prisoner dilemma. For example, [Frey and Bohnet \[1997\]](#) and [Bohnet and Frey \[1999\]](#) show that identification leads to a higher cooperation rate than anonymity when playing the prisoner dilemma, and to larger amounts offered in the dictator game. Similarly, [Charness and Gneezy \[2008\]](#) consider the effect of revealing the family name of a participant's counterpart in dictator and ultimatum games. They find that in dictator games, revealing the name of the recipient results in more generous allocations, while in the ultimatum game this has no significant effect. A number of public good experiments have also investigated the role of anonymity in giving. [Rege and Telle \[2004\]](#), for example, show that revealing the identity of the participants significantly increase contributions in a repeated public good game as identification enables people to gain social approval. Those experiments highlight how the identification of the participants positively affects interactions with other individuals. However, they do not provide insights on the effect of privacy concerns in the formation of social networks. Inspired by these experiments, we let participants identify themselves in two treatments, but we made sure that our participants did not lie when revealing their name, by checking their identity as in [Feri et al. \[2015\]](#) so as to be sure we were reliably inducing privacy concerns for all participants, and that this was common knowledge. Moreover, our participants were not forced to reveal their real name; they could choose to remain anonymous rather than publicly disclosing their name to the other members of the group. In this way we avoided a selection effect at the recruitment stage whereby we would only have been able to invite participants who stated that they were ready to reveal their name.

Beyond the privacy literature, our research is also related to the experimental literature on the provision of public goods in networks, where networks are either formed endogenously or exogenously given. There is a large amount of theoretical literature on networks and we do not attempt to summarize it in this paper – [Jackson et al. \[2008\]](#) provides a complete overview. Experimental studies with endogenous network formation mainly focus on the network architecture and its stability properties rather than on public good provision (see [Falk and Kosfeld, 2012](#); [Caldara and McBride, 2014](#)). For example, [Falk and Kosfeld \[2012\]](#) find that groups rarely converge to the theoretical equilibrium star network when exogenously given payoffs between different positions are asymmetric. Experimental studies with public good provision in networks focus on the effect of exogenously given network structures on contributions [[Laury and Holt, 2008](#), [Choi et al., 2016](#)]. For example, [Rosenkranz and Weitzel \[2012\]](#) test the predictions from the model introduced by [Bramoullé and Kranton \[2007\]](#) in the laboratory. Their treatment variable is the structure of the network. By varying the structure across treatments, they identify how individual decisions in a (local) public goods game depend on the global and local characteristics of the underlying network. They find that the ability to coordinate in terms of contributions in the public good game strongly depends on the network structure (i.e. full and star network), yet they also find that global network structure only partially determines behavior. Both investment and successful co-

ordination are significantly influenced by an agent’s local position, specifically an agent’s degree. Very few studies allow network formation and provision of public goods to be chosen endogenously as in our paper (Bravo et al., 2012; Fatas et al., 2015; Goyal et al., 2014). For example, Bravo et al. [2012] present an experimentally grounded agent-based model to investigate trust diffusion and cooperation, and demonstrate that cooperation increases and free-riders are isolated when agents can choose their links. Although there are some similarities between those studies and ours, the difference lies in the fact that we do not provide information on the past behavior of the participants as a tool for partner selection. Instead, we focus on the impact of the decision of an individual to disclose personal private information or not. Moreover, unlike under the standard approach, we ask our participants to play the contribution game before deciding which links to form. The main reason is that in our experiment, the decision to contribute is sensitive information that can be put into play later in the network formation stage. The decision to contribute or not is made without knowing what occurs later in the link formation stage, during which the participants find out which other participants decided to disclose their names. Inverting the order of decision – first linking, then contributing – would make contributions dependent on revelation by others. Yet the aim of this study is to investigate the effect of contributions – the private information that is put into play – on the decision to reveal one’s name and on the decision who to link with in the network formation stage.

Our main hypothesis is that in the absence of past interactions, an individual’s decision to reveal his/her name will be used by other participants as a predictor of the individual’s contribution, and that individuals will indeed reveal their names to signal contribution, thus influencing their choice who to establish a link with and the subsequent structure of the network. The issue is whether such a simple mechanism is sufficient to encourage more contributions, whether putting private information into play is indeed necessary and if so, why.

3 Design of the experiment

This experiment was designed to study the impact of privacy concerns as a driving force in individual network formation. The experiment consisted of two phases: the first phase was designed to create sensitive personal information, while the second phase was used to study individual linking choices. The first (contribution) phase had only a **contribution stage**. The second (network formation) phase consisted of three stages: the **signaling stage**, the **link formation stage** and the **disclosure stage**. Those two phases were repeated over 6 periods, in each of which participants were randomly matched into groups of five. Participants never met the same persons (i.e. there was *perfect stranger matching*) and never received any feedback regarding their realized payoffs in the previous periods. At the end of the experiment (i.e. after the sixth period), there was an additional stage, the **revelation and pay-off stage**, in which one out of the six periods was randomly selected for payment and the participants found out about their realized payoff and observed the information that other members of their group decided to reveal.

Table 1: EXPERIMENTAL TREATMENTS

TREATMENT	Repeated over 6 periods				DISCLOSURE STAGE
	CONTRIBUTION STAGE	SIGNALING STAGE	LINK FORMATION STAGE	REVELATION AND PAYOFF STAGE	
Baseline	Individuals decide to contribute 0 or 5		Participants decide with whom form a link. For each participant only a generic identifiers is available.		One period out of six is selected for payment. Payoff are realized and participants learn the sum of contributions in their neighborhood for that period.
Names	Individuals decide to contribute 0 or 5	Participants can decide to disclose their real name.	Participants decide with whom form a link. In addition to the generic identifier, for each participant the decision (Yes/No) to disclose their real name in the next stage is provided.	Participants' real names are revealed.	One period out of six is selected for payment. Payoff are realized. Participants real names (if disclosed) are revealed.
Names+Info	Individuals decide to contribute 0 or 5	Participants can decide to disclose their real name.	Participants decide with whom form a link. In addition to the generic identifier, for each participant the decision (Yes/No) to disclose their real name in the next stage is provided.	Participants' real names are revealed.	One period out of six is selected for payment. Payoff are realized. Participants real names (if disclosed) are revealed to the other participants along with their individual contributions.
Fake	Individuals decide to contribute 0 or 5	Participants can decide to disclose their fake name.	Participants decide with whom form a link. In addition to the generic identifier, for each participant the decision (Yes/No) to disclose their fake name in the next stage is provided.	Participants' fake names are revealed.	One period out of six is selected for payment. Payoff are realized. Participants fake names are revealed to the other participants.
Fee	Individuals decide to contribute 0 or 5	Participants can decide to pay a fee.	Participants decide with whom form a link. For each participant the decision (Yes/No) to pay the fee is provided.		One period out of six is selected for payment. Payoff are realized.

We implemented five treatments. The contribution phase was the same in all treatments (see screenshot in Figure 1), while the treatments differed in terms of the type of signal individuals could use. Table (1) gives an overview of the structure of experiment and of differences between treatments.

Only the link formation stage is of interest in the *Baseline* treatment, as no information is provided at the signaling and disclosure stages. At the linking stage, the participants simply saw a screen showing anonymous identifiers (namely 1, 2, 3, 4, and 5) of the other participants in their group, and could choose who they wished to establish a link with by ticking the relevant boxes on the screen (see Figure 2). A participant received nine ECUs for each reciprocated link with someone who contributed, while he/she paid five ECUs for each reciprocated links if he/she contributed (see below for the complete payoff structure). In this treatment, the participants could not rely on any signal to distinguish between other group members.

In the *Names* treatment, participants were asked at the signaling stage whether they wished to disclose their name to their peers at the disclosure stage – after links had been established (see screenshot *a* Figure 3). At the link-formation stage, the participants viewed a screen with anonymous identifiers of the other participants (1, 2, ...,5) and their signal: their decision (Yes or No) to disclose their name in the following disclosure stage (screenshot *b* Figure 3). Their real names were later revealed to the other participants at the disclosure stage (screenshot *c* in Figure 3). It is important to note that it is only at this stage – that is after the linking decisions – that participants could see the names of those who decided to reveal their name. In this way we are able to isolate the effect of being ready to disclose one’s name (i.e. the willingness to incur privacy costs in case of no contribution) from the effect of other characteristics possibly associated with a name that could affect the linking choice (e.g. gender, ethnicity, religion, see [Charness and Gneezy, 2008](#); [Mayer and Puller, 2008](#)). As already mentioned, it is also important to note that, in order to ensure that real names were given in the signaling stage, the experimenter checked the identity of each participant and made sure it corresponded to the name they entered. This procedure is required in order to guarantee the reliability of the private information generated in the laboratory (see also [Feri et al., 2015](#)). In order to avoid that decisions of the disclosing participants be indirectly identified by other participants through the noise they make when typing, we asked non-disclosing participants to key in a random string of characters. In this way the participants were unable to predict how many others would reveal their names when asked whether they wished to reveal their own name.³

The *Names+Info* treatment was identical to the *Names* treatment except that at the end of the experiment (i.e. in the revelation and pay-off stage) each participant was also informed about the real names of the other participants (if disclosed) as well as their individual contributions (see again Table 1). Therefore *Names* and *Names+Info* only differed in the type of information that was revealed at the very end of the experiment (see screenshot *a* and *b* Figure 4), and thus differed in

³The participants were visually isolated from one another by walls, but there was the possibility that they could hear some sounds from their neighbors.

terms of the privacy concerns they generated. The highest privacy concerns were generated by the treatment *Names+Info*, as individual contributions were identified by name, while this could only occur indirectly in the *Names* treatment (namely, if an individual established a mutual link with only one other participant). In comparison to the *Names+Info* treatment, the *Names* treatment was designed to investigate why allowing the revelation of real names helped with linking decisions: either this was because participants believed that someone who revealed his/her name was *intrinsically* more likely to have contributed, in which case results would be similar in the *Names* and *Names+Info* treatments, or participants believed that someone who revealed his/her name was more likely to have contributed because a non-contributor would not want to reveal his name. In that case, the impact of name revelation on linking decisions would be particularly pronounced in the *Names+Info* treatment, where names and contributions were linked with each other with certainty at the end of the experiment.

To check whether nicknames could have the same effect as real names, we ran a *Fake Names* treatment, which is the same as the *Names* treatment except that at the signaling stage, each participant could choose – in the disclosure stage – one of two unique nicknames (one male and one female) as provided by the experimenter. Even though individuals can freely choose their own nicknames on many online networks, we decided against allowing our participants to do this in order to control the content of the fake names. For the purpose of our experiment, it was important to have an alternative signal that did not involve privacy concerns. For this reason, as well as to ensure comparability across treatments, the experimenter checked that the chosen fake name was typed in correctly—i.e. as proposed by the experimenter—by each participant. In this treatment, we hypothesized that giving each participant a unique identity reminded them that they were dealing with other, individuated people, i.e. that their actions had an impact on individuals (see [Sanfey et al. 2003](#), [Johnson and Mislin 2011](#)).

Finally, to check for other effects of signaling mechanisms, we ran a final, *Fee* treatment, in which instead of revealing their names, the participants could decide to pay a fee in the signaling stage, in which case this signal was given at the link formation stage. At the link formation stage, the participants viewed a screen with the identifiers of the other participants along with their decisions whether to pay the fee (Yes or No) next to it. This treatment is designed to be analogous to paying a club membership fee. Unlike in the *Names* and *Names+Info* treatments, signaling that one paid the fee does not involve privacy concerns. However, like in those treatments, the participants could differentiate from others by means of a (costly) signal. Unlike in those treatments, the cost of the signal was monetary. We ran this treatment to ensure that the effect of generating a privacy concern in the *Names* and *Names+Info* treatments could not simply be explained by the willingness of those who reveal their name to incur a cost in doing so – the cost of revealing their name to others. We set the fee at a low level as other experiments have shown that the monetary equivalent of revealing one's name to other people is a low value for most people [[Grossklags and Acquisti, 2007](#), [Frik and Gaudeul, 2016](#)]. Similarly to revealing one's name, paying a fee can be viewed as showing a willingness to establish a relationship with others; thus the fee treatment enables us to

Table 2: TREATMENT AND NUMBER OF INDEPENDENT OBSERVATIONS

TREATMENT	INDIVIDUALS	INDIVIDUALS	PERIOD	INDIVIDUAL	INDEPENDENT	DYADIC
		PER GROUP	OBS	OBS	GROUP OBS	OBS
<i>Baseline</i>	25	5	6	150	30	600
<i>Names</i>	25	5	6	150	30	600
<i>Names+Info</i>	25	5	6	150	30	600
<i>Fake</i>	25	5	6	150	30	600
<i>Fee</i>	25	5	6	150	30	600

investigate this possible source of the effect of privacy decisions.

By comparing the behavior of our participants across treatments, we can isolate the pure effect of revealing one’s individual identity on the selection of partner and the formation of a social link, thus distinguishing it from the effects of other types of signals, such as giving a fake name or paying a fee. More precisely, we can identify the effect of the privacy concerns associated with being able to connect names and contributions on the probability of link formation.

The experiment was run between June and July 2013 at the laboratory of the Max Planck Institute of Economics in Jena (Germany). The experiment was programmed with the z-Tree software [Fischbacher, 2007]. An English translation of the original German instructions is provided in Appendix B. We invited 125 participants and divided them into 5 sessions. One session was held for each treatment, each one with a different pool of 25 participants playing the game for 6 periods. We paired participants in groups of five and implemented *perfect stranger matching* so that no one was in a group with the same person more than once. One period out of six was selected at random for payment. Therefore, in total we obtained $6 \times \frac{25}{5} = 30$ independent group observations per treatment (see Table 2). The participants gave written consent to participate in the study, according to the Declaration of Helsinki (Helsinki [1991]), and the procedure was carried out in compliance with the experimental rules of the Jena laboratory.⁴ The average payoff was approximately 8.90 Euro. Each session lasted for about one hour and half, and did not start until all participants were familiar with the procedure. In order to ensure that participants understood the game, a series of non-payoff relevant examples were provided in the instructions and the subjects had the opportunity to visualize the effect of various configurations of contribution and linking decisions on their payoffs (see the figure included in the translated instructions provided in Appendix B).

3.1 Matching technology and payoff structure

In each round t , at the contribution stage, each participant i decided whether or not to contribute in this round. We denote this decision as $C_t^i = \{0, 1\}$, with $C_t^i = 0$ for contribution and $C_t^i = 1$ for no contribution. Subsequently, in the signaling stage, each participant decided on a signal which differed across treatments. We denote that decision as $A_t^i = \{No, Yes\}$. For example, in *Names*, *Yes* represents the decision to reveal one’s name, while in *Fee* it represents the payment of a fee.

⁴<https://experiment.wiwi.uni-jena.de//public/rules.php>

Then, at the link formation stage, each participant i simultaneously chose with whom to link. In particular, letting I denote the set of five agents in a group, participant i made *link proposals*, a 4-tuple $p_t^i = (p_{\{I/i\}}^i) \in P_t^i = \{0, 1\}^4$ with $t = 1, 2, \dots, 6$ denoting the round. The action $p_{j,t}^i = 1$ denotes the decision by agent i to link with agent j , while $p_{j,t}^i = 0$ denotes i 's choice not to link with j in round t . A link in round t was established only if mutually agreed, that is *iff* $p_{j,t}^i \cdot p_{i,t}^j = 1$. A network is the set of all agreed upon links, $g_t = \{(i, j) \in I : p_{j,t}^i \cdot p_{i,t}^j = 1\} \in \Gamma$, where Γ is the set of all possible networks. We denote $N^i(g_t) = \{j : p_{j,t}^i \cdot p_{i,t}^j = 1, j \neq i\}$ the *neighborhood* of agent i in network g_t . This is the set of all agents with whom he/she is bilaterally linked. The number of neighbors of agent i is the cardinality of $N^i(g_t)$ and is denoted $n^i(g_t)$. The set of feasible graphs varies from the complete graph – uniform matching where each agent has $n = 4$ direct links – to the empty model where each agent has no direct link. We denote with $a_t^i = \{C_t^i, A_t^i, p_t^i\}$ participant i 's actions in round t .

The utility of agent i in any round t therefore depended on his actions and on the actions taken by his neighbors, that is

$$u_t^i(g) = \sum_{j \in N^i(g_t)} u^i(a_t^i, a_t^j) = \sum p_{j,t}^i \cdot p_{i,t}^j \cdot 9 \cdot C_t^j - 5 \cdot C_t^i \cdot \sum p_{j,t}^i \cdot p_{i,t}^j - PC_{Treat}(A_t^i, C_t^i) \quad (1)$$

where $p_{j,t}^i \cdot p_{i,t}^j \cdot 9 \cdot C_t^j - 5 \cdot C_t^i \cdot p_{j,t}^i \cdot p_{i,t}^j$ are the monetary benefits and costs related to i and j 's contributions in the contribution phase. For example, if i and j established a link in round t , and both i and j decided to contribute, the monetary benefits were 9 for both of them, while the monetary costs were 5, for a net monetary payoff for i of $9 - 5 = 4$. Those who decided not to link simply earned 0.⁵

If all individuals were forced to link with each others, the game would resemble a *linear public good game*, which can also be seen as a multi-person prisoner's dilemma.⁶ However, since the participants were free to choose their links, this contribution game is a partially excludable public good-game, as an individual could always prevent another from benefiting from his/her contribution by choosing not to link with that participant (i.e. $p_{j,t}^i \cdot p_{i,t}^j = 0$).

$PC(A_t^i, C_t^i)$ represents our manipulation variable, i.e. the costs and benefits (non-monetary in all treatments except for fee) related to i 's signaling and contribution decisions, which depend on the treatment $Treat = Baseline, \dots, Fee$ (see again Table 1 for an overview of the treatments). For example, $PC_{Names}(A_t^i, C_t^i)$ represents the non-monetary (i.e. privacy) costs and benefits in *Names*, which depends on i 's name revelation and i 's choice of contribution. In the next section, we hypothesize how the variation in these costs across treatments will affect the linking choices.

⁵In the experiment individuals could also make losses which are deducted from the show-up fee. In this case, at the end of the experiment individuals filled in a's and A's at the rate of a one-page text per 2 ECU loss (see last page example in Appendix B). This happened only twice.

⁶A multi-person prisoner's dilemma is a game in which individuals may take actions that are socially valuable but individually disadvantageous. For a linear public goods game to be a prisoners' dilemma, all of the players must be better off when they all cooperate than if they all defect, and it must be that given the actions of others, each player gets a higher payoff from defecting than from cooperating. See Bergstrom [2002].

3.2 Hypotheses

In line with the standard experimental framework for a public good game, a rational agent, who is maximizing individual utility, has no incentive to contribute in the contribution phase. In fact, each participant faces the same temptation to defect and to free-ride on the other players' contributions. However, this is inefficient as all contributing would lead to a Pareto superior outcome. Results of previous public good game experiments lead us to expect positive contributions in all treatments [Laury and Holt, 2008].

Unlike the standard procedure in experimental economics, in which anonymity is maintained among laboratory participants (see Hoffman et al., 1994), we introduced the possibility of revealing the full name of the participants and their contribution choices at the end of the experiment (*Names+Info* treatment). In this treatment, we expect that the preferences of some subjects for *social approval* (Akerlof, 1980; Rege and Telle, 2004; Charness and Gneezy, 2008) will lead to higher overall levels of contribution. The same effect could also play a role in the *Names* treatment, where overall contributions of members of one's own neighborhood g_t are revealed. Here, social approval may arise at the group level, or at the individual level if $n^i(g_t) = 1$, whereby the neighbor's contribution is known.

We denote benefits from social approval SA , with a subscript denoting the treatment. If $A_t^i = \text{Yes}$ and $C_t^i = 1$ then

$$SA_{\text{Names+Info}} > SA_{\text{Names}} > SA_{\text{Fake}} \sim SA_{\text{Fee}}$$

while if $A_t^i = \text{No}$ or $C_t^i = 0$ then there is no social approval.

Hypothesis 1: Contributions If individuals have preferences for **social approval**, we should observe a greater share of contributors in treatments in which real names are used (i.e. *Names* and *Names+Info*). In particular, we should observe higher contribution rates in *Names+Info* in which there is a certain and direct possibility to be identified as a contributor. Conversely, when there is full anonymity (i.e. *Fakes*, *Fee* and *Baseline*) this benefit does not exist and contribution levels will therefore be lower.

While subjects who contribute derive social approval from being able to associate their name with their contribution, subjects who do not contribute are subject to a *privacy concern*: they would prefer not to reveal their name. The extent of such privacy concerns can be referred to as the *privacy cost* incurred when seeing one's lack of contribution associated with one's name. These costs differ among participants. Some individuals may really dislike their name being associated with non-contributing behavior, while others may not mind. We define *privacy costs* as the costs incurred if a participant decides to reveal his/her name while he/she has not contributed. Those include the costs of social disapproval but also other potential costs associated with the leaking of negative personal information.⁷ According to this definition of privacy costs, there are no privacy

⁷In reality, there are several potential privacy costs: increased probability of being discriminated, possible identity theft, decrease in the market value of personal data, social stigma. While in this experiment the risk of identify theft is very low, we cannot exclude subsequent discrimination or social stigma. See for example Romanosky and Acquisti [2009].

costs in revealing one's name if one contributed. We rather speak of a “*taste for privacy per se*” as a driver for not revealing one's name even when one contributed.

We denote privacy costs PC , with a subscript denoting the treatment. If $A_t^i = \text{Yes}$ and $C_t^i = 0$ then

$$PC_{\text{Names+Info}} > PC_{\text{Names}} > PC_{\text{Fake}} \sim PC_{\text{Fee}}$$

while if $A_t^i = \text{Yes}$ and $C_t^i = 1$ then there are no privacy costs by this definition (although there may be costs associated with revealing one's name).

Hypothesis 2: Revelation (Privacy Concern) If *privacy costs* are on average sufficiently high, then we expect that individuals who did not contribute are less likely to signal (reveal their name) in treatments *Names+Info* and *Names*. As a result, the share of individuals choosing to signal will be lower in treatments *Names+Info* and *Names* than in other treatments, where signaling does not involve revealing one's name.

Previous experimental research has also shown that, whenever it is possible to identify a participant, the fear of shame for not contributing increases contributions more than the anticipation of social approval for contributing (see for example [Samek and Sheremeta, 2014](#)). Moreover, some individuals may have a *taste for privacy per se*, i.e. they do not wish to reveal their name even if they contributed. We can therefore further decompose the revelation behavior according to the contribution decision.

Hypothesis 3.1: Revelation & Contributions If the share of individuals that have a *taste for privacy per se* is low, while the share of participants with a preference for social approval is high, then the relative share of signaling individuals who contribute will be higher in treatments *Names* and *Names+Info*. The greatest effect will be observed in *Names+Info* where the opportunities for obtaining social approval are strongest.

Hypothesis 3.2: Revelation & No Contribution If the share of individuals that have a *taste for privacy per se* is low, while the share of participants with positive privacy costs is high, then the relative share of signaling individuals who do not contribute will be lower in treatments *Names* and *Names+Info*. The greatest effect will be observed in *Names+Info* where there is the highest risk of incurring privacy costs.

As long as privacy costs are on average sufficiently high and most participants have a relatively low taste for privacy per se, then signaling one's readiness to reveal one's name indicates that one contributed, while not revealing one's name indicates that one did not contribute. Indeed, denote $pr(C = 1)$ the expected portion of contributors and $pr(C = 1, A = \text{Yes})$ the expected portion of contributors who signal their readiness to reveal their name. Among those who are not contributors, $pr(C = 0, A = \text{Yes})$ are expected to also signal. Then the likelihood, when observing signal $A = \text{Yes}$, that the person is a contributor is

$$pr(C = 1 | A = \text{Yes}) = \frac{pr(C = 1, A = \text{Yes})}{pr(C = 1, A = \text{Yes}) + pr(C = 0, A = \text{Yes})}$$

while the likelihood when observing signal $A = No$, that the person is a contributor is

$$pr(C = 1|A = No) = \frac{pr(C = 1, A = No)}{pr(C = 1, A = No) + pr(C = 0, A = No)}$$

We hypothesize that $pr(C = 1|A = Yes) > pr(C = 1|A = No)$ in treatments *Names* and *Names+Info*. This is the case as long as $pr(C = 0, A = Yes)$ – the expected portion of individuals who reveal their name even if they did not contribute – is low, while $pr(C = 1, A = No)$ – the portion of individuals who do not reveal their name even if they contribute – is low. This is the case as long as individual generally have high privacy costs and low pure taste for privacy. Therefore, in that situation, those who revealed their own name are more likely to be a contributor than those who did not.

The possibility of giving names and identifying contributors by their readiness to disclose their name thus increases the probability of establishing a link with a contributor: individuals who disclose their name will link with one another and not with those who do not disclose their name. Since there will be more reciprocal links between contributors, overall profits will be higher in *Names+Info* than in *Names*, and higher in *Names* than in the other treatments.

Hypothesis 4 Linking Choice & Network Size. *If most participants have privacy concerns but only a few have a taste for privacy per se, then the readiness to disclose one's name is correlated with being a contributor. Anticipating this, contributors will be more likely to add those who signal than those who do not. The likelihood of a reciprocal link being established between two contributors is thus higher if they both signal than if one or the other does not. The two treatments that induce privacy concerns will therefore display more reciprocal links between contributors, and the number of such reciprocal links will be even higher in treatment Names+Info. Overall profits will thus be higher in Names+Info than in Names, and higher in these treatments than in the other treatments.*

4 Univariate analysis of the experimental data

In Table (3) we report the summary statistics (computed at individual level) of the variables of main interest. We first notice from this table that – in all treatments – the share of contributors was on average fairly low (between 15% and 27%), while the number of individuals added was quite high, approximately 3.60-3.90 out of 4 individuals (i.e. about 90% of all possible links).

In Table (4) we compare these statistics across treatments. However, in order to account for the correlation of the observations at a group-level, we rely on the mean of each statistic computed for each group and period when making comparisons across treatments. Although these treatment averages are identical to those computed at the individual level (as each mean in Table (3) is the average of the group means), the *t-test* statistics are more conservative due to the smaller number of group observations.

With regard to contributions, we observe significant differences when we compare the *Names+Info* with the *Fake* and *Fee* treatments. In particular, the share of contributors was 8% higher in *Names+Info*

Table 3: SUMMARY STATISTICS

This table reports the summary statistics computed relying on individual observations. $C_i = 1$ denotes contribution, while 0 denotes no contribution. $A_i = Yes$ denotes the choice to signal (i.e. to reveal a real/fake name, to pay a fee).

	BASELINE		NAMES		NAMES+INFO		FAKE		FEE	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Contributing Behavior Variables										
% who contribute 5 (Contributors, i.e. $C_i = 1$)	0.167	0.374	0.213	0.411	0.267	0.444	0.187	0.391	0.153	0.362
% who reveal names (or paid the fee, i.e. $A_i = Yes$)		0.447	0.447	0.499	0.48	0.501	0.88	0.326	0.547	0.499
% who reveal names & contribute (i.e. $C_i = 1, A_i = Yes$)		0.1	0.1	0.301	0.173	0.38	0.18	0.385	0.073	0.262
% who reveal names & do not contribute (i.e. $C_i = 0, A_i = Yes$)		0.347	0.347	0.477	0.307	0.462	0.7	0.459	0.473	0.501
Network Size Variables										
Nr of people added	3.82	0.625	3.58	1.051	3.607	0.882	3.773	0.625	3.707	0.719
Nr of reciprocal links	3.64	0.762	3.227	1.069	3.213	0.938	3.547	0.747	3.427	0.838
Nr of reciprocal links if $C_i = 1, A_i = Yes$			2.333	1.112	2.385	1.134	3.259	0.764	3.090	1.044
Nr of reciprocal links if $C_i = 0, A_i = Yes$			3.634	0.627	3.782	0.554	3.695	0.667	3.606	0.621
Density	0.91	0.15	0.807	0.197	0.803	0.15	0.887	0.118	0.857	0.141
Profits	2.267	9.156	2.347	9.282	2.853	9.379	2.427	9.039	1.027	6.883
N Observations	150		150		150		150		150	

Table 4: CONTRIBUTION AND NAME REVEAL: COMPARISONS ACROSS TREATMENT OF (GROUP) MEANS

This table reports in columns the mean-difference between treatments for each one of the variable listed in rows. The means are computed by keeping for each treatment the mean of each group per period (to account for group correlations), and relying on one-sided test for the equality of means in large sample. Results from pr-test are consistent. The means for each variable (computed at individual level) are reported in Table (3).

	NAMES VS BASELINE	NAMES VS FAKE	NAMES VS FEE	NAMES+INFO VS NAMES	NAMES+INFO VS FAKE	NAMES+INFO VS FEE
Contributing Behavior Variables						
% who contribute 5 (Contributors)	0.047 (0.146)	0.027 (0.282)	0.06* (0.092)	0.053 (0.127)	0.08** (0.048)	0.113*** (0.008)
% who reveal names (or paid the fee)		-0.433*** (0.000)	-0.1** (0.028)	0.033 (0.258)	-0.4*** (0.000)	-0.067 (0.110)
% who reveal names & contribute		-0.08** (0.024)	0.027 (0.008)	0.073*** (0.000)	0.033 (0.231)	0.100*** (0.000)
% who reveal names & do not contribute		-0.353*** (0.000)	-0.127*** (0.009)	-0.04 (0.215)	-0.393*** (0.000)	-0.167*** (0.001)
Network Size Variables						
Nr of people added	-0.24*** (0.009)	-0.167** (0.011)	-0.127 (0.101)	0.027 (0.606)	-0.193** (0.021)	0.1 (0.103)
Nr of reciprocal links	-0.413** (0.014)	-0.32** (0.033)	-0.2 (0.135)	-0.013 (0.471)	-0.333** (0.011)	-0.213* (0.083)
Nr of reciprocal links if $C_i = 1, A_i = 1$		-0.976*** (0.002)	-0.831** (0.040)	-0.008 (0.492)	-0.968*** (0.001)	-0.822** (0.028)
Nr of reciprocal links if $C_i = 0, A_i = 1$		-0.249* (0.076)	-0.108 (0.278)	0.018 (0.463)	-0.231* (0.053)	-0.090 (0.280)
Density	-0.103** (0.014)	-0.08** (0.032)	-0.05 (0.135)	-0.003 (0.471)	-0.083** (0.011)	-0.053* (0.084)
Profits	0.08 (0.445)	-0.08 (0.443)	1.32*** (0.005)	0.507 (0.162)	0.427 (0.232)	1.827*** (0.000)
N	60	60	60	60	60	60

*p-values in parentheses: *p < 0.10, **p < 0.05, ***p < 0.01*

than in *Fake* (p -value=0.048), and 11% higher in *Names+Info* than in *Fee* (p -value=0.008). When we compare *Names* with the other treatments, the share of contributors was higher in *Names* but not statistically significantly so.

Result 1 *The share of contributors was higher in the treatments where individuals could reveal their name (i.e. Names and Names+Info), though only significantly so in Names+Info, where direct social approval was possible. However, the share of contributors was fairly low in all treatments (below 30%).*

Moreover, the share of individuals who revealed fake names in *Fake* was substantially higher than the share of individuals who revealed their real name in *Name+Info* (+40%, p -value=0.000) and in *Names* (+43%, p -value=0.000). The share of individuals who paid the fee in *Fee* was also higher than the share of those revealing their name in *Names* and *Names+Info*, although significantly so only in comparison with *Names* (+10%, p -value=0.028)—see again Table (4).

Result 2 *The share of individuals choosing to signal was lower in treatments where a privacy concern could arise.*

Those first two results show that our participants were sensitive to using their real name in the laboratory. In particular, they appear to have been particularly sensitive to the revelation of names along with their contribution in *Names+Info*.

We can further decompose the statistics for individuals who chose to signal as a function of their contribution. As Table (4) suggests, even though there are only few striking differences between treatments when comparing the share of individuals who signal *and* contribute, there are substantial differences when comparing the share of those who signal *and do not* contribute. In line with the hypothesis of higher aversion toward social disapproval (shame), the share of individuals who signaled (reveal their name) and did not contribute was significantly higher in *Fake* than in both *Names* (35%, p -value=0.000) and *Names+Info* (39%, p -value=0.000). Similarly, the share of individuals who signaled (paid the fee) and did not contribute was significantly higher in *Fee* than the share of individual who signaled (revealed their name) and did not contribute in both *Names* (12.7%, p -value=0.009) and *Names+Info* (16.7%, p -value=0.001). As highlighted above, we also observe a greater share of individuals who contributed and revealed their name in *Names+Info* than in *Names* (+7%, p -value=0.000) and *Fee* (+10%, p -value=0.000).

Result 3 *The share of individuals who chose to signal and contribute was significantly higher in Names+Info than in Names or Fee (but not compared to Fake). In contrast, the share of individuals who chose to signal and not contribute was generally significantly lower when privacy concerns could arise (i.e. in both Names+Info and Names).*

Those results suggest that individuals did care about social approval, but they were mostly concerned about their real name being directly (*Names+Info*) or indirectly (*Names*) associated with non-contribution at the end of the experiment. That is, privacy costs mattered more than social approval.

In Table (4) we finally compare the size of the networks across treatments. To study this issue, we consider two indicators of network size: the *number of people added*, which provides us with the

potential size of the network, and the *number of reciprocal links* (i.e. the degree of a node), which give us a measure of the actual size of the network. We also construct a measure of the density of the network, which gives us an index of the degree of dyadic connection in a population (i.e. a normalized version of the average degree of the network). An ideal, fully connected network would have a density of 1.⁸

We observe that the number of people added was significantly lower in *Names* than in *Baseline* (-0.24 individuals, p -value=0.009) and lower in *Names* than in *Fake* (-0.167 individuals, p -value=0.021), while there are no significant differences when comparing *Names+Info* and *Fee*. A similar pattern can be observed when looking at the number of reciprocal links in the *Names+Info* treatment: when privacy concerns arise, individuals tend to add less people and obtain fewer reciprocal links. We provide a detailed overview of the decomposition of contributions and people added in Figures (A.1)-(A.5) in the Appendix, where we focus on individuals who contributed and distinguish those who revealed or did not reveal their name.⁹

Moreover, we see that contributors who revealed their name had substantially smaller networks. For example, the number of reciprocal links was significantly lower both in *Names+Info vs Fake* (-0.968 individuals, p -value=0.001) and in *Names vs Fake* (-0.976, p -value=0.0031).¹⁰ As a result, when real names were used instead of fake names, the density of the network was significantly lower (between 8% and 10%). Similar numbers are observed when we compare those treatments with *Fee*, although no significant differences emerge in terms of network density. We do not find significant differences between *Names+Info* and *Names*.

Finally, we observe that profits tended to be higher in *Names+Info* than in all other treatments, although only significantly so when compared to *Fee* (1.827 €, p -value=0.000).

Result 4 *When real names were used, the size of the network (i.e. the number of people added and of reciprocal links) was lower. In particular, the largest decrease was observed in Names+Info and for contributors who revealed their name. In this case, on average, the number of reciprocal link was lower by one compared with other treatments.*

Result 5 *Profits tended to be higher when privacy concerns were strongest (i.e. Names+Info) but only significantly so in comparison with Fee.*

Overall, Result 1-5 suggest that, when privacy concerns arose, the actual and potential size of the network was smaller, and fewer individuals chose to signal (reveal their name) without contribut-

⁸The calculation is straightforward - known connections divided by maximum possible connections. For binary data, as in our case, the density is simply the proportion of all of the possible dyadic connections which are actually present, that is the ratio of the number of adjacencies that are present divided by the number of unique pairs (i.e. $(n \times (n - 1))/2$).

⁹We find that non-contributors almost always add everyone in their group. From these figures, we observe that those who revealed their name are wary of participants who did not reveal their name in *Names+Info*, while in *Fee* the contributors who did not pay a fee are actually wary of individuals who paid a fee. In *Fake*, almost all participants “reveal” a fake name, so the treatment is very similar to *Baseline* where there are no names to help one to choose among the various links.

¹⁰This is the only case in which the averages computed at the individual level differ from those computed at the group level. In this case indeed, the number of individuals per group varies and depends on the number of individuals who decided to disclose.

ing. As a result, profits tended to be higher (although not always significantly so). Those results suggest that the combination of the revelation of contributions at the end of the experiment along with real names led the participants to select fewer but more profitable links. As supplementary evidence, we consider the network graphs across treatments: when real names were used, we seldom observe the uniform matching that often emerged in the other treatments (see [online Appendix](#)).

5 Multivariate analysis of reciprocal links: Dyadic regressions

In this section, we run a multivariate analysis of the probability of establishing a reciprocal link between two individuals. This approach enables us to jointly consider both contribution and disclosure as determinants of the formation of a link, as well as other individual attributes. A number of methods have been developed to study network formation empirically (see [Bramoullé and Fortin \[2010\]](#)). We rely here on dyadic regressions as in [Fafchamps and Gubert \[2007a,b\]](#) [Mayer and Puller \[2008\]](#), in which each observation expresses a relationship between pairs of individuals. Therefore, any pair of individuals i and j , either linked or not, constitutes an observation. The advantage of this method is that it uses information from both individuals (i.e. g_{ij} and g_{ji}) to identify the impact of individual similarities or dissimilarities (e.g. in the disclosing behavior) on the probability of establishing a link. Therefore, in order to apply this approach, the data must be re-organized in order to express the set of all possible links (except with oneself) for each individual. In our case, this consists in 4 observations for each individual (i.e. a total of $6 \cdot 4 \cdot 25 = 600$ observations per treatment, see [Table \(2\)](#)). Formally, we express the propensity of establishing a link with $g_{ij}^* = \alpha + \beta X_{ij} + u_{ij}$, and g_{ij} denoting the existence of a link between individuals i and j such that

$$g_{ij} = \begin{cases} 1 & \text{if } g_{ij}^* > 0 \\ 0 & \text{otherwise} \end{cases}$$

g_{ij} is an $N \cdot (N - 1)$ matrix, X_{ij} is a series of $N \cdot (N - 1)$ matrices.¹¹

From [Equation 1](#), we can see that a participant i derives a utility $u_i(g)$ from his network g , and this utility depends on the network structure. In order to tie back dyadic regressions with individual decisions, we need to assume *i*) the separability of the utility function [1](#) and *ii*) symmetry. That is, $u_i(g) = \sum_j u_i(g_{ij})$ and $u_i(g_{ij} = 1) - u_i(g_{ij} = 0) = u_j(g_{ij} = 1) - u_j(g_{ij} = 0)$. The default assumption, which is consistent with our experiment ($g_{ij} = g_{ji}$ for all i and j as we have unidirectional dyadic relationships), is that every individual in a group is a potential partner.

The main idea behind this approach is that individuals can choose which links to form with one another by performing a cost-benefit analysis of each link with the following form:

¹¹The total number of possible pairs is N^2 but we dropped the N_{ii} pairs on the diagonal. This resulted in 600 (=25*24) observations for each treatment.

$$g_{ij} = \begin{cases} 1 & \text{if } B(C^j, d_{ij}) - \text{Cost}(C^i, d_{ij}) + e_{ij} > 0 \\ 0 & \text{otherwise} \end{cases}$$

where g_{ij} denotes the existence of a reciprocal link between individuals i and j , d_{ij} is a set of characteristics of individual i and j , and e_{ij} is the residual effect. In the experiment, $B(C^j, d_{ij})$ includes the monetary benefits of a *reciprocal* links with an individual j . $B(C^j, d_{ij})$ is always equal to 9 if subject j contributed (i.e. $B(C_t^j = 1, \cdot) = 9$), and equal to 0 if he/she did not contribute ($B(C^j = 0, \cdot) = 0$). In *Names* and *Names+Info*, $B(C^j, d_{ij})$ also includes benefits from social approval for individuals who have contributed and revealed their name. $\text{Cost}(C^i, d_{ij})$ represents both the monetary and non-monetary costs of establishing a reciprocal link. The monetary cost of a link is always equal to 5 if the individual contributed and 0 if not. In *Names* and *Names+Info*, individuals may also incur a privacy cost, depending both on the type of action associated with the links (e.g. contribution) and on the sensitivity of the subject to the privacy concern associated with this information. In the *Fee* treatment there is an additional monetary cost of 1.¹²

As [Fafchamps and Gubert \[2007a\]](#) highlight, there is an identification problem, which is related to the way in which regressors X_{ij} enter the regressions. When using dyadic data, we can distinguish two type of regressors: attribute w_{ij} of the link (such as geographical distance), and attributes x_i and x_j of the nodes i and j . Regressors must enter the dyadic regression in a symmetric fashion so that the effect of (x_i, x_j) on g_{ij} is the same of (x_j, x_i) on g_{ji} , and thus $\beta X_{ij} = \beta X_{ji}$. In our setting, we assume

$$\text{Prob}(g_{ij} = 1) = \Lambda(\alpha + I(x_{1i} \neq x_{1j})\beta_1 + I(x_{2i} \neq x_{2j})\beta_2 + w_i + u_{ij})$$

where x_1 is a vector of dummies: $x^{\text{Both Reveal}}$, $x^{\text{Both Contribute}}$, x_2 is a vector of dummies: $x^{\text{One Reveal}}$, $x^{\text{One Contribute}}$, and $\Lambda(\cdot)$ is the logistic cumulative distribution. For each couple i and j , only two variables can be simultaneously activated, being mutually exclusive. For example, $x^{\text{Both Contribute}}$ and $x^{\text{One Reveal}}$.

Therefore, β_1 can be interpreted as the vector coefficients of the effects of equal choices (e.g. both revealing), while β_2 is the effect of different choices (e.g. only one reveals the name). In other words, the effects of the combinations (1,1) is β_1 , while β_2 gives the effects of combinations $\{(1,0) (0,1)\}$. The base category is the effect of the combination (0,0).

It is important to note that only the signals are observable to the other participants. The variable w_i represents the number of people added in each round by individual i and aims to control for the individual propensity to establish a link. In alternative specifications, we also controlled for individual risk aversion and general trust.¹³

¹²There are also typing costs, which are constant across participants and subjects as even not-disclosing participants needed to type.

¹³To derive the level of individual risk aversion we used the tasks of [Heinemann et al. \[2009\]](#), while the level of general trust was measured in the exit questionnaire with general questions about trusting others as in [Gächter et al. \[2004\]](#).

Finally, as it is essential to correct standard errors in dyadic regressions, we assume that $E(u_{ij}, u_{i,k}) \neq 0$ and $E(u_{ij}, u_{k,j}) \neq 0$ for all k . We therefore cluster standard error terms to account for correlation within dyads (i.e. the standard error term is also dyadic).

5.1 Multivariate analyses: results

Logit estimates of the probabilities that a link exists in a pair (i.e. the results from the dyadic regressions) are presented in Table (5), where the coefficients represent marginal effects estimated for each of the five treatments.¹⁴ Before comparing the results across treatments (which is complicated by an identification problem inherent in logit regression, see Long [2009]), let us start by discussing the general results across all treatments.

Table 5: PROBABILITY OF ESTABLISHING A RECIPROCAL LINK

The dependent variable is a dummy equal to 1 if a link between any pair of individuals i and j exists, and zero otherwise. The data are therefore organized to express for each individual the set of all possible links (except with oneself). This consists of 4 observations for each individual per period (i.e. a total of $600=25*4*6$ observations per treatment). *One contribute* is a dummy variable equal to 1 if either i or j has contributed and zero otherwise, *Both contribute* is a dummy variable equal to 1 if both i and j has contributed and zero otherwise, *One reveal* is a dummy variable is equal to 1 if either i or j has chosen to signal and zero otherwise, *Both reveal* is a dummy variable equal to 1 if both i and j has contributed and zero otherwise, while *People Added* is the total number of individual added per round. Standard errors are clustered at dyad level.

	BASELINE	NAMES	NAMES+INFO	FAKE	FEE
<i>One contribute (i.e. $C_i \neq C_j$)</i>	-0.0346 (0.024)	-0.0497 (0.039)	-0.1785*** (0.043)	-0.0883*** (0.029)	-0.2161*** (0.056)
<i>Both contribute (i.e. $C_i = C_j = 1$)</i>	-0.0266 (0.060)	-0.2327** (0.098)	-0.2469*** (0.076)	-0.0634 (0.080)	-0.0514 (0.064)
<i>One reveal (i.e. $A_i \neq A_j$)</i>		-0.0484 (0.037)	-0.0532 (0.055)	0.0796 (0.075)	-0.0390 (0.051)
<i>Both reveal (i.e. $A_i = A_j = 1$)</i>		0.1726*** (0.033)	0.1986*** (0.045)	0.3043 (0.191)	0.0752 (0.048)
<i>People added</i>	0.0992*** (0.014)	0.1234*** (0.017)	0.1275*** (0.010)	0.1067*** (0.010)	0.0801*** (0.017)
Log-likelihood	-131	-185	-199	-156	-156
N	600	600	600	600	600

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

As expected, in all treatments, the higher the number of people added, the higher the probability of establishing a reciprocal link. However, when real names were used (i.e. in *Names* and in *Names+Info*), the probability of establishing a link for an extra individual added was higher (about 12% vs 10%) than in other treatments. If we consider that the average number of people added is

¹⁴This is equivalent to estimating a model with full interactions between each variable and dummies for treatments.

lower in this treatment (see Table 3), this result – as in the univariate analysis – indicates that the participants took more care in choosing their links.

The probability of establishing a reciprocal link between two individuals who contributed (i.e. $C_i = C_j = 1$) was generally lower than the probability of establishing a reciprocal link between two individuals who did not contribute and did not signal (i.e. the base category), but it was statistically significantly lower only for *Names* and *Names+Info*. For example, in *Names+Info* the probability that two contributors established a reciprocal link was 25% lower (significant at 5% level) than between two non-contributors. In this treatment, this probability was also lower even when only one of the two individuals had contributed (i.e. $C_i \neq C_j$). As shown in the univariate analysis, the share of contributors was low (about 30%). Moreover, contributors had smaller networks especially when real names could be revealed. Therefore, a non-contributors was more likely to establish a link with a non-contributor than with a contributor.

However, if both individuals performed the same action, that is they chose to signal ($A_i = A_j = 1$), the probability of establishing a link was significantly higher when real names were used than when other communication signals were used. For example, in *Names* the probability that a reciprocal link was established between two individuals who revealed their name was about 17% higher than the probability of establishing a link between two individuals who did not disclose their name, while in *Names+Info* it was about 20% higher. The effects of asymmetric actions (i.e. $A_i \neq A_j$) are various and milder, but in all cases they are never significant. These results suggest that the mutual revelation of names was a crucial determinant for establishing a link in these treatments.¹⁵

We now compare the results in *Names* and *Names+Info* with those obtained in the other treatments. However, as the traditional tests of the equality of coefficients across groups in non-linear regressions are complicated by identification problems (i.e. the residual variation confound the magnitude of the effects, see Long, 2009), we compare predicted probabilities across groups at multiple levels of the variable *People Added*. The results from this exercise are reported in Figure 5, which shows the differences in probability between treatments along with the confidence intervals. As these figures highlight, the differences are significant (i.e. the confidence interval does not contain the zero) if we compare the predicted probabilities of *Names* vs *Baseline*, *Names* vs *Fake* and *Names+Info* vs *Fake*, when only few people are added (less than 3). For example, when only few individuals were added in *Names*, the probability of establishing a reciprocal link was 20-30% higher than in *Fake*. No significant differences across treatment emerge when many individuals were added. This result is in line with the univariate analysis, suggesting that in these treatments the individuals who used real names tended to select fewer but more specific links.

Result 5 *The probability of establishing a reciprocal link increased significantly when both individuals*

¹⁵We refrain from adding any interaction terms in this regression as the model is non linear and the coefficients on the interaction terms (i.e. how the effect of one variable changes when another variable changes) do not provide the change in the partial effect of the variables on the conditional mean function. In addition, in some cases, the results of hypothesis tests are an artifact of the functional form and do not necessarily have an economically meaningful content. See Greene [2010].

revealed their name (in Names and Names+Info). Moreover, when name revelation was possible, the probability of establishing a link between two non-contributors that did not signal was significantly higher than the probability of establishing a link between two contributors.

Table 6: REGRESSIONS OF INDIVIDUAL PROFITS ACROSS TREATMENTS

The dependent variable is the total amount of individual profits resulting from all the established links in each period. Contribution dummy is a dummy equal to 1 if the individual has contributed, while Reveal dummy is a dummy equal to one if the individual chose to signal. Standard errors are clustered at individual level.

	BASELINE	NAMES	NAMES+INFO	FAKE	FEE
<i>Contribution dummy</i>	-11.0303*** (1.188)	-13.1274*** (3.906)	-15.0668*** (4.840)	-20.0326*** (3.200)	-8.1129*** (2.443)
<i>Reveal (paid fee) dummy</i>		1.4390 (1.677)	4.9413** (2.272)	1.0582 (3.168)	-2.4126 (1.679)
<i>Reveal · Contribution</i>		-2.3469 (3.734)	-0.1026 (5.178)	5.5787 (3.622)	0.2503 (4.662)
<i>Constant</i>	4.1051*** (0.198)	4.7391*** (0.782)	4.5171*** (1.231)	4.2307 (2.838)	3.5712*** (1.004)
	-487	-470	-480	-488	-459
Individual fixed effects	Yes	Yes	Yes	Yes	Yes
N Observations	150	150	150	150	150

*p<0.10, ** p<0.05, ***p<0.01

Finally, Table (6) reports the determinants of individual profits per period. The dependent variable is now the total amount of individual profits resulting from all the established links in each period. Those results suggest that, across all treatments, contributors earned significantly less than non-contributors (i.e. the base category). This is remarkable in *Fake*, where contributors earned about 20 ECUs less than non-contributors. This result is consistent with previous experimental research on public-good games.

However, we observe that in *Names+Info*, individuals who disclosed their name earned about 5 ECUs more (significant at the 5% level) than those who did not disclose their name. This means that in this treatment, an individual who committed to disclose his name at the end of the experiment was able to earn 5 ECUs more than an individual who did not commit to disclose his name. However, since the interaction term between dummies for contribution and disclosure behavior is not significant, we can also infer that an individual who contributed and committed to reveal his/her name did not earn significantly more than a non-contributor who committed to disclose his/her name. Therefore, although two contributors managed to establish a reciprocal link by both revealing their names, there were also a number of non-contributors who managed to establish a reciprocal link with a contributor by disclosing their name. The results (not reported but available upon request) that control for the number of reciprocal link established are consistent.

Result 6 *In all treatments contributors earned significantly less than non-contributors. However, when high privacy concerns arose (i.e. in Names+Info) individuals who disclosed their personal names earned significantly more (if non-contributors), or lost significantly less (if contributors), than non-disclosing in-*

dividuals.

In line with the univariate analysis outlined above, we therefore conclude that individuals were able to select more valuable links through the endogenous choice of using their real name. However, this privacy mechanism was not sufficient to exclude non-contributors with low privacy costs, who were thus able to exploit contributors by revealing their name. The simple mechanism of endogenous choice of revealing one's name was however effective as it allowed contributors to obtain higher profits as long as they revealed their name. Introducing the possibility to incur privacy costs made it possible for contributors to avoid linking with the type of individuals who dare misbehave (not contribute) only under the veil of anonymity.

5.2 Robustness checks

To check the robustness of the previous results, we consider two other procedures that are generally used in network analysis for studying dyadic data to account for correlation across unobservables: the quadratic assignment procedure (QAP) and the multilevel analytic approach.

The QAP procedure consists in scrambling the dependent variable in the regression (in our case the dependent variable measuring the reciprocal link), while keeping the independent variable in the original observation positions. After scrambling several times, an empirical sampling distribution is obtained. When comparing the actual coefficients with the empirical distribution, no significant relationship is expected between the dependent and independent variables under the null of no-statistical association between the variables.

The multilevel analytic approach (Snijders et al., 1995; Snijders and Bosker, 2011) distinguishes between two levels of data in personal network studies: the level of the relation (the dyad) is the first and lowest level (level 1) and includes attributes of dyads and attributes of alters within these dyads; the level of individual (level 2) includes attributes of the ego (such as personal characteristics). In our case, we estimate a three-level random-intercept logit model with an additional group level (level 3) as our individuals are randomly matched into groups.

The results (not reported but available upon request) obtained from both procedures are consistent with results from dyadic regressions, and robust also to the inclusion of individual measures of risk aversion (see also Rosenkranz and Weitzel, 2012) and general trust.

6 Discussion & conclusion

There are many types of social networks, each with different goals, but a common and primary objective is to foster connections between users. These connections may be important for facilitating other tasks, such as finding jobs (e.g. LinkedIn), making friends (e.g. LiveJournal), dating (e.g. Meetic) or sharing news (e.g. Twitter). However, connecting with other users may involve incurring costs, such as the cost of building and maintaining relationships, but also costs resulting from the misuse by others of personal shared information [Krasnova et al., 2010]. Therefore, as

in real life, individuals need to form impressions of others in order to decide whether to establish (or continue) a relationship. However, especially in online environments, traditional identity cues that are normally available, such as accent or style of dress, are not present. Other types of information may be used, but whether they are available is mainly under the control of social network managers. The amount and the type of information that can be shared is therefore an essential part of the formation of relationships [Tidwell and Walther, 2002]. For example, the elements shared in the profile may act as a signal to communicate the personal quality of the user (Lampe et al. [2007]).

In this paper we studied how the possibility of revealing one's name affected the selection of partners when forming individual networks. Our experiment consisted in both a "contribution phase" and a "network phase". In the contribution game, the individuals decided whether or not to contribute, while in the "network game" they had to decide with whom to establish links. Contributions were exchanged only if reciprocal links were established.

We exogenously varied across treatments the type of signals individuals could use to identify contributors, as well as to signal personal contribution. More precisely, while the matching technology remained constant, treatments differed in terms of the costs associated with the use of the signaling opportunity. In two treatments, the participants could decide whether to reveal their real name, and their real name was either associated with their contribution at the end of the experiment (*Names+Info* treatment), or not (*Names* treatment). Therefore, *privacy costs* could arise for those individuals who decided to form links without contributing as they could be directly or indirectly identified at the end of the experiment. Conversely, individuals could obtain *social approval* by contributing. In the three other treatments, there was either no signaling possibility (*Baseline* treatment), or participants could disclose a fake name (*Fake* treatment) or they could pay a fee (*Fee* treatment). Those treatments were designed to explore the ways, other than social approval and privacy concerns, in which revealing one's name could affect social network formation. In the *Names* treatment, social approval and privacy concerns were reduced compared with the *Names+Info* treatment, so the main effect there was whether those who reveal their name were perceived as more likely to have contributed in the absence of the privacy costs or benefits associated with being able to connect (non-)contribution with individual names at the end of the experiment. In the *Fake* treatment, we examined how revealing one's name could act as a way of enhancing the feeling that one is dealing with distinguishable individuals, rather than a group of anonymous participants. In the *Fee* treatment, we analyzed the possible effect of revealing one's name as signaling one's readiness to incur some type of cost (monetary in *Fee*, in terms of loss of privacy in *Names* and *Names+Info*) to participate in the game.

In line with previous experiments on public-good games, our results indicate that participants cared about social approval and tended to contribute more in treatments when they could be directly identified as contributors (i.e. in *Names+Info*). Moreover, participants considered the revelation of information about their non-contributing behavior to be sensitive: they were less willing to reveal their name when they did not contribute. Finally, we observed that the probability of

establishing a link was significantly higher when both individuals signaled that they would reveal their real names than when other signaling opportunities were used. As a result, we also observed that when real names could be used, the size of individual networks was substantially smaller while average profits were higher, especially for those who disclosed their name. The decision to reveal their name therefore helped participants to find suitable matches (i.e. contributors) among the numerous linking opportunities. However, a number of individuals with low privacy concerns also gave their real names even when not contributing. We do not have insights into their intentions, but it is likely that they did so with the conscious objective of passing as a likely contributor.

Our experiment indicates that giving individuals the possibility to reveal their name along with their contribution generates higher levels of contribution, but more importantly it significantly reduces the number of individuals who do not contribute and signal. Higher contributions may arise due to the desire to obtain social approval – which is obtained if one also reveals one’s name. They may also arise because of the expectation by contributing individuals that they will be better able to attract productive links later on by revealing their name. The logic is that by revealing one’s name, one attracts contributors, but revealing one’s name without contributing also generates privacy costs. Therefore, individuals who want to attract contributors but avoid privacy costs have to contribute in addition to revealing their name.

Comparison with the other treatments shows that a system of endogenous name revelation and contribution works mainly thanks to the fear of incurring privacy costs. Indeed, contributions decrease hand-in-hand with privacy concerns, from the *Names+Info* treatment to the *Names* treatment. Contributions are lower in the *Names* treatment where they cannot generally be directly associated with individuals, so the effect of revealing names is not mainly due to a belief that those who reveal their name are intrinsically more likely to be contributors. Contributions are also lower if only fake names are used, so the effect of revealing names is not mainly due to the fostering of the identification of other group participants as individuals. Similarly, contributions are lower if a system of costly signaling (*Fee*) is used, so the effect of revealing names is also not due to the willingness to incur some type of cost (loss of privacy) in order to be considered as a worthy person to link with.

Why did real names act as a relatively reliable signal in a context in which deception is so beneficial? What made a real name a more reliable signal of contribution than a fake name? In order to be a reliable signal, the disclosure of the name must be correlated with the qualities the partner should have: on seeing this signal one should be confident that the participant is a contributor. Yet, for this to occur, the signal must be particularly costly for non-contributors. Non-contributing individuals must therefore bear high privacy costs. In this case participants feel that they can “trust” the signal; they are confident that those who reveal their name are more likely to be contributors. In our setting, participants could not build a reputation as they only played the game once with the same group of people. Therefore, the reputation that was at stake was their own personal reputation in other contexts, in the event that they met other participants in their group again outside

of the laboratory. When repeated interactions are possible, it is likely that fake names would work as well as real names to identify contributors, as individuals build reputations attached to their fake name. This simple experiment thus suggests that privacy concerns can be used to isolate free-riders even in a context without repeated interactions.

The results obtained from our simple (partially excludable) public-good and network formation experiment lead us to conclude that privacy mechanisms (such as revealing one's name) significantly affect the selection of partners and the consequent structure of the network. Allowing participants to lift the veil of anonymity, thereby increasing their privacy costs, can help mitigate the exploitative behavior of some network members and promote the establishment of fewer but more valuable links between people.

References

- A. Acquisti and J. Grossklags. Privacy and rationality in individual decision making. *Security & Privacy, IEEE*, 3(1):26–33, 2005. 5
- Alessandro Acquisti, Curtis R Taylor, and Liad Wagman. The economics of privacy. *Journal of Economic Literature*, 54(2), 2016. 2
- George A Akerlof. A theory of social custom, of which unemployment may be one consequence. *The Quarterly Journal Of Economics*, 94(4):749–775, 1980. 13
- Robert Axelrod, Ross A Hammond, and Alan Grafen. Altruism via kin-selection strategies that rely on arbitrary tags with which they coevolve. *Evolution*, 58(8):1833–1838, 2004. 4
- Michael Bacharach and Oliver Board. The Quality of Information in Electronic Groups. *Netnomics*, 4(1):73–97, 2002. ISSN 1385-9587, 1573-7071. doi: 10.1023/A:1014970712376. URL <http://link.springer.com/article/10.1023/A:1014970712376>. 2
- Russell W Belk. Extended self in a digital world. *Journal of Consumer Research*, 40(3):477–500, 2013. 2
- Theodore C Bergstrom. Evolution of social behavior: individual and group selection. *The Journal of Economic Perspectives*, 16(2):67–88, 2002. 12
- I. Bohnet and B.S. Frey. The sound of silence in prisoner’s dilemma and dictator games. *Journal of Economic Behavior & Organization*, 38(1):43–57, 1999. 6
- Yann Bramoullé and Bernard Fortin. Social networks: Econometrics. *The New Palgrave Dictionary of Economics*, 3, 2010. Basingstoke: Palgrave Macmillan. 20
- Yann Bramoullé and Rachel Kranton. Public goods in networks. *Journal of Economic Theory*, 135(1): 478–494, 2007. 6
- Giangiaco Bravo, Flaminio Squazzoni, and Riccardo Boero. Trust and partner selection in social networks: An experimentally grounded model. *Social Networks*, 34(4):481–492, 2012. 4, 7
- Michael Caldara and Michael McBride. An experimental study of network formation with limited observation. Working Papers 141501, University of California-Irvine, Department of Economics, 2014. 6
- G. Charness and U. Gneezy. What’s in a name? Anonymity and social distance in dictator and ultimatum games. *Journal of Economic Behavior & Organization*, 68(1):29–35, 2008. 6, 9, 13
- Syngoo Choi, Edoardo Gallo, and Shachar Kariv. Networks in the laboratory. In Yann Bramoullé, Andrea Galeotti, and Brian Rogers, editors, *The Oxford Handbook of the Economics of Networks*. Oxford University Press, 2016. 3, 6

- Marcel Fafchamps and Flore Gubert. The formation of risk sharing networks. *Journal of Development Economics*, 83(2):326–350, 2007a. [20](#), [21](#)
- Marcel Fafchamps and Flore Gubert. Risk sharing and network formation. *The American economic review*, 97(2):75–79, 2007b. [20](#)
- Armin Falk and Michael Kosfeld. It’s all about connections: Evidence on network formation. *Review of Network Economics*, 11(3):1–36, 2012. [6](#)
- Enrique Fatas, Miguel A Meléndez-Jiménez, Antonio J Morales, and Hector Solaz. Public goods and decay in networks. *SERIEs*, 6(1):73–90, 2015. [7](#)
- F. Feri, C. Giannetti, and N. Jentzsch. Disclosure of Personal Information under Risk of Privacy Shocks. *Journal of Economic Behavior and Organization*, *Forthcoming*, 2015. [6](#), [9](#)
- U. Fischbacher. z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10:171–178, 2007. [11](#)
- B.S. Frey and I. Bohnet. Identification in democratic society. *Journal of Socio-Economics*, 26(1):25–38, 1997. [6](#)
- Eric J. Friedman and Paul Resnick. The Social Cost of Cheap Pseudonyms. *Journal of Economics & Management Strategy*, 10(2):173–199, June 2001. ISSN 1530-9134. doi: 10.1111/j.1430-9134.2001.00173.x. URL <http://onlinelibrary.wiley.com/doi/10.1111/j.1430-9134.2001.00173.x/abstract>. [2](#)
- Alisa Frik and Alexia Gaudeul. The relation between privacy protection and risk attitudes, with a new experimental method to elicit the implicit monetary value of privacy. Center for European, Governance and Economic Development Research Discussion Paper 296, University of Goettingen, Department of Economics, 2016. URL <https://ideas.repec.org/p/zbw/cegedp/296.html>. [10](#)
- Simon Gächter, Benedikt Herrmann, and Christian Thöni. Trust, voluntary cooperation, and socio-economic background: survey and experimental evidence. *Journal of Economic Behavior & Organization*, 55(4):505–531, 2004. [21](#)
- Avi Goldfarb and Catherine E Tucker. Privacy regulation and online advertising. *Management Science*, 57(1):57–71, 2011. [5](#)
- Sanjeev Goyal, Stephanie Rosenkranz, Utz Weitzel, and Vincent Buskens. Individual Search and Social Networks. Working Papers 2014.49, Fondazione Eni Enrico Mattei, April 2014. [3](#), [7](#)
- Bryan S Graham. Methods of identification in social networks. Technical report, National Bureau of Economic Research, 2014. [3](#)

- W. Greene. Testing hypotheses about interaction terms in nonlinear models. *Economics Letters*, 107(2):291–296, 2010. [23](#)
- Jens Grossklags and Alessandro Acquisti. When 25 cents is too much: An experiment on willingness-to-sell and willingness-to-protect personal information. In *Proceedings of the Sixth Workshop on the Economics of Information Security (WEIS'07)*, pages 7–18, 2007. [5](#), [10](#)
- Frank Heinemann, Rosemarie Nagel, and Peter Ockenfels. Measuring strategic uncertainty in coordination games. *The Review of Economic Studies*, 76(1):181–221, 2009. [21](#)
- DO Helsinki. Statements from the vancouver group. *British Medical Journal*, 302:1194, 1991. [11](#)
- Elizabeth Hoffman, Kevin McCabe, Keith Shachat, and Vernon Smith. Preferences, property rights, and anonymity in bargaining games. *Games and Economic Behavior*, 7(3):346–380, 1994. [13](#)
- Matthew O Jackson and Asher Wolinsky. A strategic model of social and economic networks. *Journal of Economic Theory*, 71(1):44–74, 1996. [3](#)
- Matthew O Jackson et al. *Social and economic networks*, volume 3. Princeton university press Princeton, 2008. [6](#)
- Noel D Johnson and Alexandra A Mislin. Trust games: A meta-analysis. *Journal of Economic Psychology*, 32(5):865–889, 2011. [10](#)
- Hanna Krasnova, Sarah Spiekermann, Ksenia Koroleva, and Thomas Hildebrand. Online social networks: why we disclose. *Journal of Information Technology*, 25(2):109–125, 2010. [25](#)
- Cliff AC Lampe, Nicole Ellison, and Charles Steinfield. A familiar face (book): profile elements as signals in an online social network. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 435–444. ACM, 2007. [26](#)
- Susan K Laury and Charles A Holt. Voluntary provision of public goods: experimental results with interior nash equilibria. *Handbook of experimental economics results*, 1:792–801, 2008. [6](#), [13](#)
- J Scott Long. Group comparisons in logit and probit using predicted probabilities. Working Paper, Department of Sociology, University of Indiana, 2009. [22](#), [23](#)
- Adalbert Mayer and Steven L Puller. The old boy (and girl) network: Social network formation on university campuses. *Journal of Public Economics*, 92(1):329–347, 2008. [9](#), [20](#)
- Plato. *The Republic*. 380 BCE. Book 2, Chapter 3, B. Jowett Trans. [2](#)
- Mari Rege and Kjetil Telle. The impact of social approval and framing on cooperation in public good situations. *Journal of Public Economics*, 88(7):1625–1644, 2004. [6](#), [13](#)

- Sasha Romanosky and Alessandro Acquisti. Privacy costs and personal data protection: Economic and legal perspectives. *Berkeley Technology Law Journal*, 24(3):1061–1101, 2009. [13](#)
- Stephanie Rosenkranz and Utz Weitzel. Network structure and strategic investments: An experimental analysis. *Games and Economic Behavior*, 75(2):898–920, 2012. [6](#), [25](#)
- Anya Savikhin Samek and Roman M Sheremeta. Recognizing contributors: an experiment on public goods. *Experimental Economics*, 17(4):673–690, 2014. [14](#)
- Alan G Sanfey, James K Rilling, Jessica A Aronson, Leigh E Nystrom, and Jonathan D Cohen. The neural basis of economic decision-making in the ultimatum game. *Science*, 300(5626):1755–1758, 2003. [10](#)
- Tom Snijders, Marinus Spleen, and Ronald Zwaagstra. The use of multilevel modeling for analysing personal networks: Networks of cocaine users in an urban area. *Journal of Quantitative Anthropology*, 5(2):85–105, 1995. [25](#)
- Tom AB Snijders and Roel Bosker. *Multilevel Analysis: An Introduction To Basic And Advanced Multilevel Modeling*. SAGE Publications Ltd, 2011. [25](#)
- John Suler. The online disinhibition effect. *Cyberpsychology & behavior*, 7(3):321–326, 2004. [2](#)
- Lisa Collins Tidwell and Joseph B Walther. Computer-mediated communication effects on disclosure, impressions, and interpersonal evaluations: Getting to know one another a bit at a time. *Human communication research*, 28(3):317–348, 2002. [26](#)

Figure 1: CONTRIBUTION STAGE

COMPUTER SCREEN AT THE CONTRIBUTION STAGE

You can invest either 0 or 5 ECU in this round of the experiment

If you invest 0 ECU, then people with whom you establish reciprocal links will receive 0 ECU and you will pay nothing.

If you invest 5 ECU, then people with whom you establish reciprocal links will receive $1.8 * 5$ ECU and you will pay 5 ECU per reciprocal links.

How many ECU do you want to invest? 0 ECU
 5 ECU

Your contribution will not be shown to other participants at the end of the experiment.

Figure 2: BASELINE TREATMENT

COMPUTER SCREEN AT THE LINKING STAGE

Please make your choice

	Your choice
Group Member 1	<input type="checkbox"/> Add
Group Member 2	<input type="checkbox"/> Add
Group Member 3	<input type="checkbox"/> Add
Group Member 4	<input type="checkbox"/> Add
Yourself	

Figure 3: NAMES TREATMENT

a.

COMPUTER SCREEN AT THE SIGNALING STAGE

Would you like to give your name to be shown to others? Yes
 No

Your name will be deleted from our records after the experiment.

b.

COMPUTER SCREEN AT THE LINKING STAGE

Please make your choice

	Will show name?	Your choice
Group Member 1	No	<input type="checkbox"/> Add
Group Member 2	No	<input type="checkbox"/> Add
Group Member 3	Yes	<input type="checkbox"/> Add
Group Member 4	Yes	<input type="checkbox"/> Add
Yourself	Yes	

Your contribution will not be shown to other participants at the end of the experiment.

c.

COMPUTER SCREEN AT THE DISCLOSURE STAGE

	Name	Your choice
Group Member 1		Added
Group Member 2		Not added
Group Member 3	Alexia Gaudeul	Added
Group Member 4	Caterina Giannetti	Added
Yourself	Harold Selbstlos	

Your contribution will not be shown to other participants at the end of the experiment.

Figure 4: NAME AND NAME+INFO TREATMENT: DIFFERENCE AT REVELATION STAGE

a.

COMPUTER SCREEN AT THE REVELATION AND PAYOFF-STAGE

The following table shows names and contributions of the group members and who you added in the period chosen for payment

	Name	Your choice
Group Member 1		Added
Group Member 2		Not added
Group Member 3	Alexia Gaudeul	Added
Group Member 4	Caterina Giannetti	Added
Yourself	Harold Selbstlos	

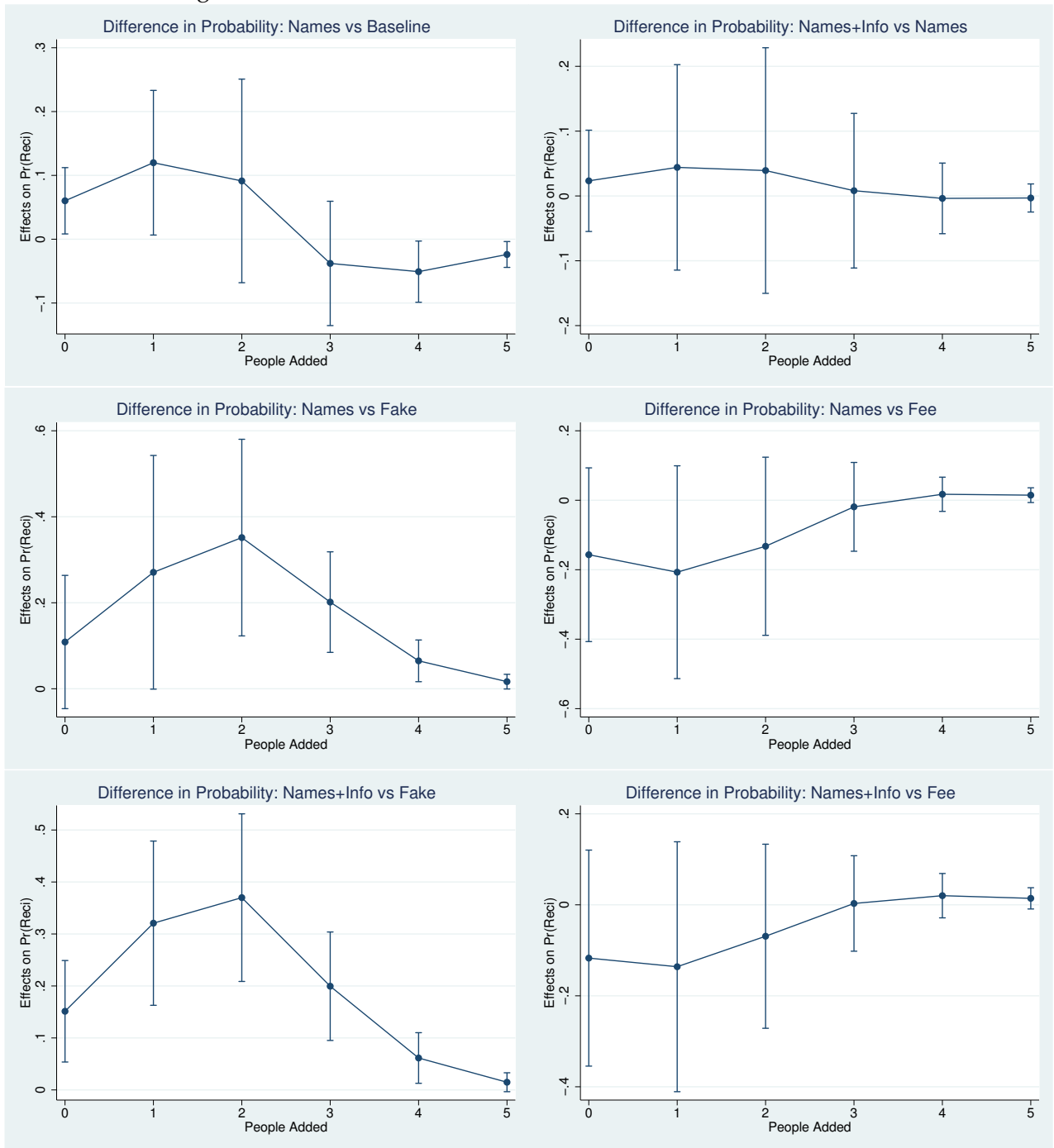
b.

COMPUTER SCREEN AT THE REVELATION AND PAYOFF-STAGE

The following table shows names and contributions of the group members and who you added in the period chosen for payment

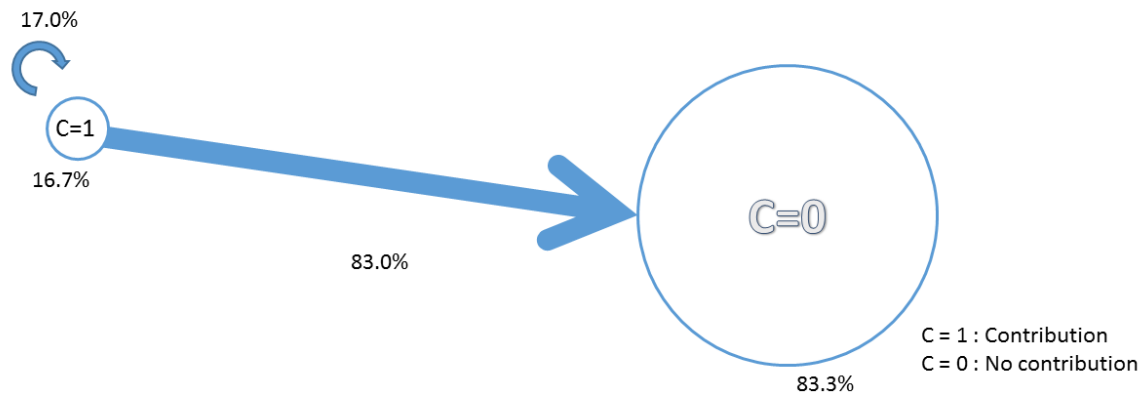
	Contribution	Name	Your choice
Group Member 1	0 ECU		Added
Group Member 2	5 ECU		Not added
Group Member 3	0 ECU	Alexia Gaudeul	Added
Group Member 4	5 ECU	Caterina Giannetti	Added
Yourself	5 ECU	Harold Selbstlos	

Figure 5: DIFFERENCE IN PROBABILITY OF ESTABLISHING A LINK



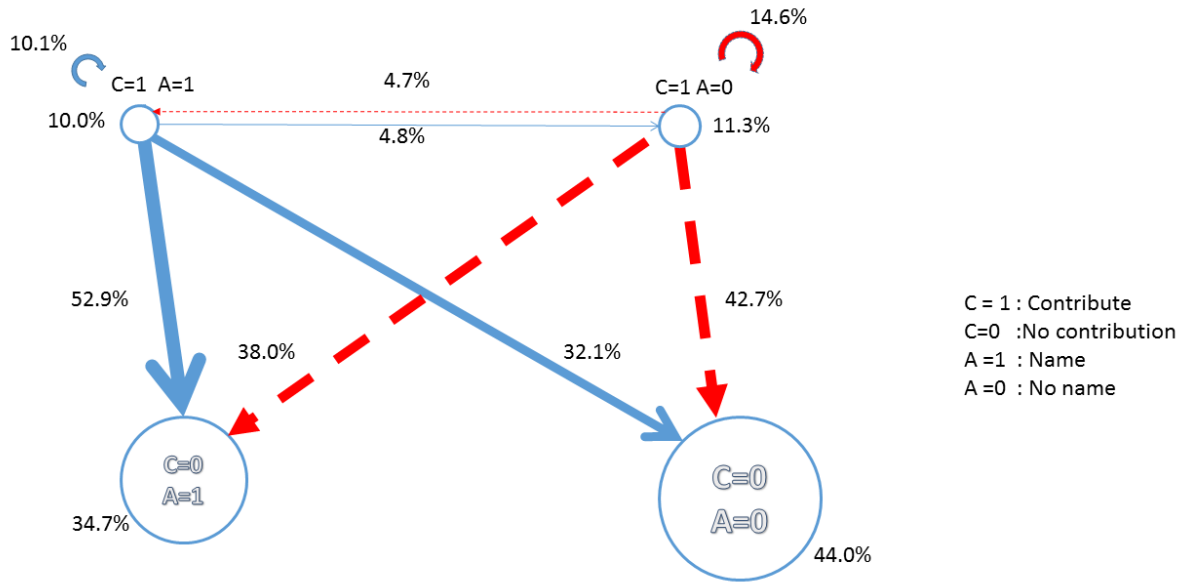
A Graphical representation of links and contributions, by treatment

Figure A.1: BASELINE



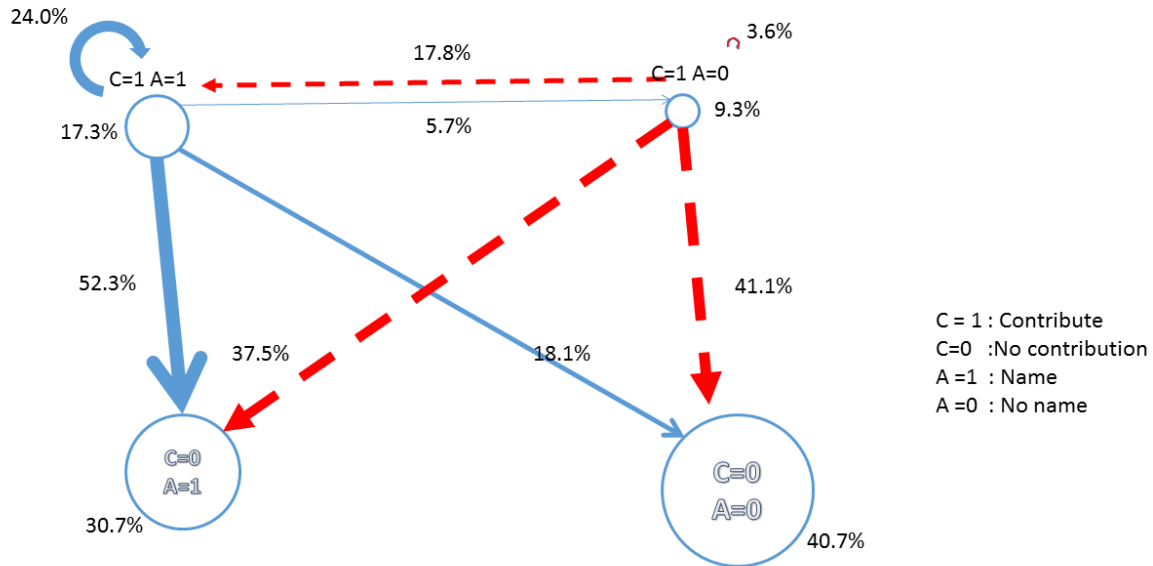
The share beside each circle represents the share of individuals in each category. For example, 16.7% of the individuals contributed (i.e. $C=1$) in the public good game. The share beside each blue arrow represents the share of individuals added by individuals who contributed (i.e. $C=1$) out of all people added. They sum up to 100%. For example, 17.0% of people added by contributors were individuals who also contributed.

Figure A.2: NAMES



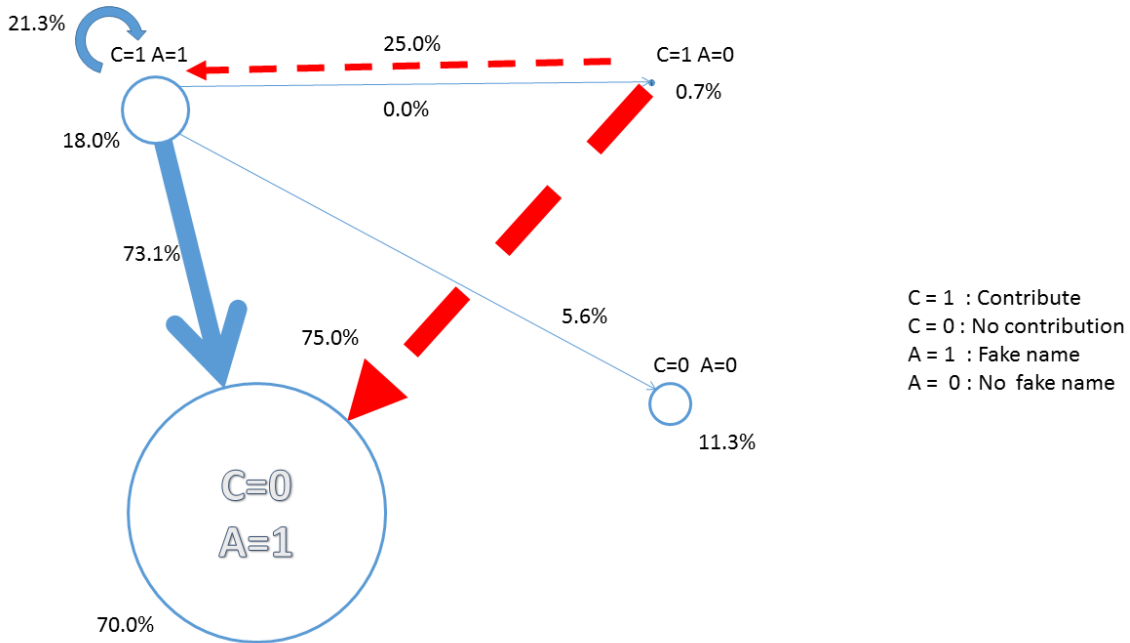
The share beside each circle represents the share of individuals in each category. For example, 10.0% of the individuals contributed in the public good game and revealed the name (i.e. C=1 and A=1). The share beside each blue arrow represents the share of individuals added by individuals who contributed and revealed the name (i.e. C=1 and A=1) out of all people added. They sum up to 100%. For example, 52.9% of people added by contributors who revealed their name were individuals who revealed their name and did not contribute. To know the total share of individuals added who revealed their name, both shares of individuals who revealed the name (disregarding contribution) must be added (i.e. 52.9% + 10.1%). Similarly, the shares beside each red-dashed arrow represents the share of individuals added by individuals who contributed and *did not* reveal their name (i.e. C=1 and A=0) out of all people added.

Figure A.3: NAMES+INFO



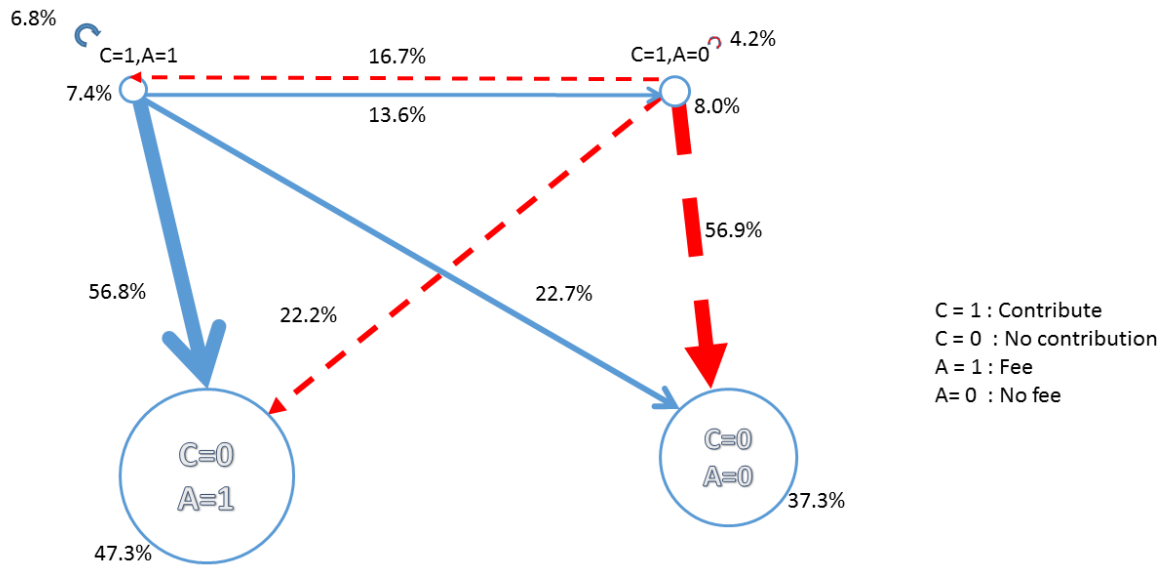
The share beside each circle represents the share of individuals in each category. For example, 9.3% of the individuals contributed in the public good game and did not reveal the name (i.e. C=1 and A=0). The share beside each blue arrow represents the share of individuals added by individuals who contributed and revealed the name (i.e. C=1 and A=1) out of all people added. They sum up to 100%. For example, 18.1% of people added by contributors who revealed their name were individuals who did not reveal their name and did not contribute. To know the total share of individuals added who revealed their name, both shares of individuals who revealed the name (disregarding contribution) must be added (i.e. 52.3+24.0%). Similarly, the shares beside each red-dashed arrow represents the share of individuals added by individuals who contributed and *did not* reveal their name (i.e. C=1 and A=0) out of all people added.

Figure A.4: FAKE NAMES



The share beside each circle represents the share of individuals in each category. For example, 70.0% of the individuals did not contribute in the public good game and revealed the name (i.e. C=0 and A=1). The share beside each blue arrow represents the share of individuals added by individuals who contributed and revealed their name (i.e. C=1 and A=1) out of all people added. They sum up to 100%. For example, 5.6% of people added by contributors who revealed their name were individuals who did not reveal their name and did not contribute. To know the total share of individuals added who revealed their name, both shares of individuals who revealed their name (disregarding contribution) must be added (i.e. 73.1+ 21.3%). Similarly, the shares beside each red-dashed arrow represents the share of individuals added by individuals who contributed and *did not* reveal their name (i.e. C=1 and A=0) out of all people added.

Figure A.5: FEE



The share beside each circle represents the share of individuals in each category. For example, 37.3% of the individuals did not contribute in the public good game and did not pay the fee (i.e. $C=0$ and $A=0$). The share beside each blue arrow represents the share of individuals added by individuals who contributed and paid the fee (i.e. $C=1$ and $A=1$) out of all people added. They sum up to 100%. For example, 13.6% of people added by contributors who paid the fee were individuals who did not pay the fee and contributed. To know the total share of individuals added who paid the fee, both shares of individuals who paid the fee (disregarding contribution) must be added (i.e. $56.8\% + 6.8\%$). Similarly, the shares beside each red-dashed arrow represents the share of individuals added by individuals who contributed and *did not* pay the fee (i.e. $C=1$ and $A=0$) out of all people added.

B Instructions

Treatment Name+Info

INSTRUCTIONS

Welcome and thank you for your participation! You will earn an amount of money that depends on your decisions and the decisions of the other participants in this experiment. It is therefore very important that you thoroughly read these instructions.

If you have a question, please raise your hand. We will then come to you and answer your question. If you violate this rule, we will be forced to exclude you from the experiment.

Please turn off your mobile phone now!

You will make decisions in the course of the experiment. All results of the study will be kept strictly confidential.

Your earnings will be calculated in ECU (Experimental Currency Units). 1 ECU corresponds to €0.50. At the end of today's session, your total earnings will be calculated and converted into euros to be paid to you confidentially and in cash.

While very unlikely, you may make losses in this experiment. If so, you will be asked to fill the a's and A's of a page of text for every 2 ECU of losses, up to a maximum of three pages. An example of the text to be filled is to be found on the last page of the instructions.

After having read the instructions, you will be given the opportunity to answer a few control questions to check your understanding of the experiment. The experiment begins when all participants have answered the control questions correctly. In the course of the experiment, you will have to make several decisions that affect you and other participants. Once the experiment is complete, we will ask you to complete a questionnaire.

EXPLANATION OF THE EXPERIMENT

As a starting amount, you receive 10 ECU. You will first be randomly assigned to a group with four other participants. Here you have to make three consecutive decisions.

DECISION 1:

First, you must decide how many ECU you want to invest. You can choose between investing either 0 or 5 ECU.

When you invest 0 ECU, every person with whom you will have formed reciprocal links in a later stage of the experiment will obtain 0 ECU and you will pay nothing. When you invest 5 ECU, every person with whom you formed reciprocal links will obtain 9 ECU and you will pay 5 ECU for each of those mutual connections.

All other members of your group also decide independently and without knowing your decision how much they want to invest.

DECISION 2:

In this stage, you need to decide if you want to give out your name to other members of your group. If you choose to do so, then please write your full name in the field provided. Then wait for an experimenter to come to you and compare the name you gave with the name on your ID card.

Please put your ID card on the table to the right of the computer keyboard with the face up if you decided to disclose your name, and face down if you chose not to disclose the name. Please do not talk to the experimenter about your decision when he / she comes to check your credentials. He / she will then enter a code so that you can move to the next decision in the experiment.

Your decision whether to disclose your name or not will be communicated to the other group members in the next stage. Your name itself will be shown to them only after they decided whether to add you to their personal list or not.

Your name will be deleted from our electronic record after the experiment, so that the experimenter will only know if you gave your name or not, but will not know your name.

DECISION 3:

In this stage, you see a list of the other members of your group, shown as "group member 1", "group member 2", "group member 3" and "Group Member 4". Each group member sees a different order of the group members, so that the person who appears on your personal list as "group member 1" may be displayed in the list of another person as "group member 3".

No one in your group can see how much you invested in decision 1. You also cannot see how much other group members have invested.

In addition to this, you can see next to the identifier of a group member whether he / she has decided to make his / her real name known. Other group members can also see if you decided in decision 2 to give out your real name.

You need to decide which members of the group you want to add your personal list, by making a check mark in the box next to the group member. If you change your mind, you can also remove the check mark.

Once you made your final decision, please click OK. All other members of your group also decide who they wish to add their personal list. The group members do not see at this stage who added who to his / her personal list.

Once you made decision 3, you will see the screen again together with your decision who you want to add and the names of those people who decided to give them. The other participants in your group also see your real name if you decided to do so (otherwise not). Please click OK when you are ready to proceed.

You will then be assigned to a new group and repeat the same steps as above. You will be assigned to 6 groups in total. You will never be more than once with the same person in the same group. In other words, the members of each group to which you are assigned are people with whom you were not in a group previously in this experiment.

Before your earnings are calculated for this experiment, we will ask you a series of questions regarding your decisions during the experiment, as well as some questions about yourself.

DETERMINATION OF PAYOUTS

At the end of the experiment the computer will randomly select one of the groups you were assigned to and you will be paid according to the results of your decisions and the decisions of the other participants in this group. Other participants in this group also get payoffs from their decisions in this group.

You do not know beforehand which group will be selected to determine the payout. You should therefore make your decisions in all six groups carefully since any of them could decide the amount of your payment at the end of the experiment. There is no connection between the groups since each group is composed of new participants and you do not know the results of the previous group.

Once a group is selected for payment, the program displays both the information that you had available at the time of decision 3, as well as who you added to your personal list and how much you invested in this group. Note that this time you not only see the names of people who decided to reveal them, but also, how much each group member invested. Likewise, the other group members will learn your real name (if you have decided in decision 2 to do so) as well as how much you invested.

You will also learn how many mutual connections you made. This is determined as follows: The program compares the people on your personal list to their own personal list:

- If a person on your personal list also has you on their own personal list, a mutual connection is made between the two of you.
In this case, you receive the amount that the person invested in Decision 1 (0 or 5 ECU) multiplied by 1.8 (so you get either 0 or 9 ECU). The other person receives the amount you invested in Decision 1 (0 or 5 ECU) multiplied by 1.8 (ie, they receive either 0 or 9 ECU). You also pay the cost of your investment (0 or 5 ECU) for each mutual connection.
- If either you did not list the other person or if the other person did not list you, then no connection is made.
In this case, you get nothing from the other and the other gets nothing from you, and you do not pay anything.

Overall, you therefore get, for each group member with whom you established a mutual connection, their investment multiplied by 1.8 minus your investment. In addition, you get 5 ECU as a show up fee and ECU 10 as a starting amount.

SUMMARY:

- You will be assigned to six groups over the experiment. Each group is composed of different people. For each group, you will need to make a few decisions in succession.
- At the end, we will ask you to fill out a questionnaire about yourself.
- Once all decision-making situations are over and you have answered all questions, the computer determines which group will determine your payoff. Then, your payoff in this group is determined and displayed along with the details of its calculation. Other members of the group also get their payoff from this group.

CONTROL QUESTIONS

First, you are asked to try out how different combinations of your decisions and those of other group members affect your earnings. In the following exercise, you are asked to decide how much you want to invest and how much you think other participants invest. You are also asked who you would add to your list and whether they would add you on their personal list.

	Contribution	Your choice	Choice by other	Net Payoff
Group Member 1	<input type="radio"/> 0 ECU <input checked="" type="radio"/> 5 ECU	<input checked="" type="checkbox"/> Add	<input checked="" type="checkbox"/> Add	$5 * 1.8 - 5$
Group Member 2	<input type="radio"/> 0 ECU <input checked="" type="radio"/> 5 ECU	<input checked="" type="checkbox"/> Add	<input type="checkbox"/> Add	0
Group Member 3	<input checked="" type="radio"/> 0 ECU <input type="radio"/> 5 ECU	<input checked="" type="checkbox"/> Add	<input checked="" type="checkbox"/> Add	$0 * 1.8 - 5$
Group Member 4	<input type="radio"/> 0 ECU <input checked="" type="radio"/> 5 ECU	<input checked="" type="checkbox"/> Add	<input checked="" type="checkbox"/> Add	$5 * 1.8 - 5$
Yourself	<input type="radio"/> 0 ECU <input checked="" type="radio"/> 5 ECU			

Participation fee + Endowment + Earnings from links – Cost of links

= Total Payoff

$$5 + 10 + (2 * 5 * 1.8) - (3 * 5)$$

= 18 ECU

Example (with reference to graph above): You established a mutual connection with GM 1, GM 3 and GM 4 but not with GM 2, who did not add you back. You invested 5 ECU. Two of the individuals with whom you made a mutual link invested 5 ECU while one of them invested 0 ECU. You therefore get $2 * 9 = 18$ ECU and pay $3 * 5$ ECU = 15 ECU. Your total earnings for this part of the experiment consists of 5 ECU (participation fee), 10 ECU (endowment) + 18 ECU (links from the other group members) - 15 ECU (your contribution multiplied by the number of mutual links established). This means you would earn in this case a total of 18 ECU.

You are invited to try out at least 4 different scenarios by changing some parameters in your decisions and those of others before you can proceed to answer some questions to check your understanding of the experiment. The experiment will continue only once all participants answered all control questions correctly.

Control questions:

1. Assume that you invested 0 ECU and established a mutual relation with two other participant, only one of whom invested 5 ECU. What is your total payoff?
 - a. $5+10+9-0=24$ ECU
2. Assume that you invested 5 ECU and established a mutual relation with three other participant, only two of whom invested 5 ECU. What is your total payoff?
 - a. $5+10+2*9-3*5=18$ ECU
3. Assume that you invested 5 ECU and established a mutual relation with three other participant, none of whom invested 5 ECU. What is your total payoff?
 - a. $5+10+0-3*5=0$ ECU
4. Assume that you invested 0 ECU and established a mutual relation with two other participant, both of whom invested 5 ECU. What is your total payoff?
 - a. $5+10+2*9-0=33$ ECU

EXAMPLE OF THE TASK FOR EACH EURO LOSS:

Please fill out all A's and a's with a pen.

Ea quae est secuta hieme, qui fuit annus Cn. Pompeio, M. Crasso consulibus, Usipetes Germani et item Tencteri magna [cum] multitudine hominum flumen Rhenum transierunt, non longe a mari, Rhenus influit quo. Causa fuit quod transeundi from Suebis complures annos exagitati bello premebantur et Agri Cultura prohibebantur. Sueborum gens est longe maxima et bellicosissima Germanorum omnium. Hi centum pagos habere dicuntur, ex quibus quotannis Singula milia armatorum bellandi causa educunt ex finibus. Reliqui, qui domi manserunt, se atque illos alunt; hi rursus in vicem anno post in armis sunt, illi domi-retentive. Sic neque nec Agri Cultura ratio atque usus belli intermittitur. Sed privatized ac separati agri apud eos nihil est, neque longius remanere anno uno in loco colendi causa licet. Neque frumento multum, sed partem maximam Lacte atque pecore vivunt multum sunt in venationibus; quae res et cibi genere et cotidiana exercitatione et libertate vitae, quod a nullo Pueris officio aut disciplina adsuefacti nihil omnino faciunt contra voluntatem, et vires alit et immani corporum magnitudine homines efficit. Atque in eam se consuetudinem adduxerunt ut neque locis frigidissimis vestitus praeter pelles habeant quicquam, quarum propter exiguitatem magna est pars corporis aperta et laventur in fluminibus.

Mercatoribus aditus est magis eo ut quae bello ceperint quibus vendant habeant, quam quo ullam rem ad se importari desiderent. Quin etiam iumentis, quibus maxime Galli delectantur quaeque impenso parant PRETIO, Germani importatis non utuntur, sed quae sunt apud eos nata, parva atque deformia, haec cotidiana exercitatione summi ut sint laboris efficiunt. Equestribus proeliis Saepe ex equis desiliunt ac pedibus proeliantur, equos Same date remanere Vestigio adsuefecerunt, ad quos se celeriter, cum usus est, recipiunt: neque eorum moribus turpius quicquam aut inertius habetur quam ephippiis uti. Itaque ad quemvis numerum ephippiatorum equitum quamvis pauci adire audent. Vinum omnino ad se non importari patiuntur, quod ea re ad laborem referendum remollescere homines atque effeminari arbitrantur.

"De Bello Gallico", Julius Caesar, Book IV