The Formation of Job Referral Networks Evidence from a Lab-in-the-Field Experiment in Urban Ethiopia[*]

A. Stefano Caria[†] Ibrahim Worku [‡]

August 8, 2013

[*]We are indebted to Marcel Fafchamps, Bart Minten, and Alemayhu Seyoum Taffesse for their support during this project. We would like to thank Johannes Abeler, Abigail Barr, Marcel Fafchamps, Edoardo Gallo, Megenagna Gashaw Guade, Derek Headey, Jeremy Magruder, Bart Minten, Esteban Ortiz Ospina, Simon Quinn, Pieter Seernels, Alemayehu Seyoum Taffesse, and seminar participants at the Ethiopian Development Research Institute (Addis Ababa), Centre for Experimental Social Sciences (Oxford), and the 2013 conference of the Centre for the Study of African Economies for useful comments and suggestions. Our gratitude also goes to the Ethiopian Development Research Institute, the Kombolcha City administration and officials, and the enumerators and supervisors whose contributions made this study possible. We finally would like to thank the participants and respondents for their goodwill and patience. All mistakes remain naturally ours. We acknowledge financial support from the International Food Research Policy Institute. Stefano Caria further acknowledges financial support from the Economic and Social Research Council, UK.

[†]DPhil Candidate, Centre for the Study of African Economies, Department of Economics, University of Oxford. Email stefano.caria@economics.ox.ac.uk

[‡]Researcher, Ethiopia Support Strategy Program II, IFPRI

Abstract

Exclusion from job contact networks constitutes a major disadvantage for labour market participants in settings where referral hiring is common and information about jobs hard to obtain. In a mid-size town in northern Ethiopia, where these mechanisms are at work, we observe that many individuals do not access local job contact networks. Using a model of strategic network formation and insights from behavioural decision theory, we hypothesize that workers would integrate poorly connected peers both when this maximises their chances of referral and when self-regarding motives are removed, because of other-regarding concerns. We devise an experimental design adapted from Beaman and Magruder (2012) to test these hypotheses. Our results lend broad support to the assumption of strategic network formation. In a setting where competition for referrals makes connections with central players undesirable, agents tie preferably to peripheral peers. However, in treatments where other regarding considerations are made salient, players do not seem to choose links with peripheral peers more often than at random. Our findings suggest that a modification of field incentives, for example via a reform in referral hiring procedures, can generate more inclusive job contact networks.

1 Introduction

Social networks often help individuals find employment by providing job referrals and information about job vacancies (Granovetter, 1995; Ioannides and Loury, 2004; Topa, 2011).[1] Exclusion from information and referral networks thus constitutes a major disadvantage for labour market participants, with profound consequence for economic equity (Calvo-Armengol, 2004; Calvo-Armengol and Jackson, 2004). In developing countries, the welfare consequences of scarce access to information and referrals are frequently aggravated by the lack of efficient matching mechanisms, high job search costs, and absent or insufficient unemployment insurance (J-PAL, 2013). Network formation processes can reduce or reinforce these initial disadvantages. Individuals' network formation decisions have been shown to be influenced by a host of social and economic incentives (Fafchamps and Moradi, 2009; Beaman and Magruder, 2012; Beaman et al., 2013). It is unclear however whether individuals respond to the incentives that are created by the structure of the network and hence whether they systematically integrate or exclude poorly connected peers (Calvo-Armengol, 2004). Innovative policy tools to address exclusion in labour markets can take advantage of an understanding of these underlying dynamics of network formation. Interest for such novel intervention designs is currently high (African Development Bank, 2012; J-PAL, 2013; World Bank, 2013).

This paper experimentally investigates whether individuals create connections with poorly integrated peers in two different economic environments. In the first environment competition for scarce referrals makes it in the agent's interest to link to peripheral peers. In the second environment self-regarding concerns are removed and other-regarding considerations come to the fore. Empirical analyses of social networks typically reveal an unequal distribution of links across individuals (Jackson, 2008). The job contact networks we document in the observational data collected as part of our study confirm this pattern.[2] Indeed the distribution

[1] Researchers have collected a large body of descriptive and econometric evidence on these mechanisms. See Topa (2011) for a recent review. In Ethiopia, referral hirings are common (Mano et al., 2010), and the use of social networks in job search is widespread and has a positive impact on the exit rate from unemployment (Serneels, 2007). The data we collected in a middle-size urban town in northern Ethiopia confirms the importance of social interactions: 41 percent of employed individuals have heard of their current occupation through their social contacts and 29 percent have received an explicit referral.

[2] See figures 2 and 3 in the appendix. In this figures, an undirected job contact link is said to exist if i has reported any exchange of information or referrals from i to j, or from j to i. In Figure 2 we depict the thus defined links of one of the block-level social networks in our sample. Blue squares represent individuals. Lines represent the job contact links. The isolated squares on the left of the

of links in local networks resembles the power distribution generated by models of preferential attachment. In these models, individuals are more likely to link with currently central peers and initial advantages reinforce over time.[3] Yet, if networks are formed strategically, peripheral individuals would be integrated in job contact networks whenever better connected peers have the right incentives to tie to them. Calvo-Armengol (2004) proposes a model where competition for referrals makes links to peripheral partners valuable. In such environment, asymmetric network architectures cannot be sustained in equilibrium (Calvo-Armengol, 2004; Calvo-Armengol and Zenou, 2005). This predictions are reinforced if individuals choose links also on the basis of two widespread other-regarding preferences: social welfare maximisation and inequality aversion (Fehr and Schmidt, 1999; Charness and Rabin, 2002).

We hypothesize that individuals understand and respond to the incentives created by the structure of the network and hence that they will link to peripheral peers when this maximises their chances of referral. These assumptions underly models of strategic network formation such as (Calvo-Armengol, 2004). We further hypothesize that individuals care about the effects of network structure on peer welfare. When self-regarding concerns are removed, individuals will still choose to connect to peripheral peers if this maximizes social welfare and outcome equality.[4] We devise an experimental design adapted from Beaman and Magruder (2012) to test these hypotheses. A common problem in analyses of network formation is that network variables may be correlated with unobserved characteristics of individuals. Our question is thus best tackled in an experimental setting where initial networks can be imposed exogenously. Furthermore, in analyses of observational data it may be hard to disentangle the different layers of social interaction. Conditions created in the lab enable us to focus the analysis on job referral networks alone. In the lab we also have control over incentives, which we can make clear and salient. This enables us to "give the best shot" to the model of incentive-sensitive network formation and hence develop a credible first test. Lastly, we strengthen the external

picture represent the individuals with no block-level job contact links. We define the number of job contact links an individual has as his or her network degree. In figure 3 we plot the distribution of the network degree of the respondents. The modal degree is zero, but a number of individuals have relatively large networks. More information about the data used to produce these figures is given in section 4

[3]For models of preferential attachment see Jackson (2008). Central individuals may be better placed to collect information about vacancies. Furthermore, theoretical models stress that central individuals are more valuable partners in risk sharing arrangements (Bramoulle and Kranton, 2007) and can be trusted more because of higher social collateral (Karlan et al., 2009).

[4]If it did only of these two things, we would expect significant heterogeneity in behaviour.

validity of the experimental design by introducing field context in the subject pool and in the task to be carried out. Following the taxonomy of Harrison and List (2004) our study can be defined as an *artefactual field experiment*.

We run our experiment with young adult dwellers of randomly sampled blocks in a small town in Ethiopia characterized by an expanding formal sector and an extensive use of job contact networks. Some of the participants in each lab session are randomly drawn to carry out a small remunerated task in the lab and to subsequently make a referral for the same task. Participants are assigned positions in a pre-determined, undirected, irregular friendship network and, if given the job, one of their ties is selected at random to receive the referral. Before the game is played, individuals are informed of the whole structure of the network and have the opportunity to indicate two further agents with whom they would like to be linked. Our analysis focuses on this link formation decision. The linking preferences of a single, randomly drawn "unemployed" participant are activated unilaterally. This network formation rule switches off other-regarding concerns, as participants' choices have no effect on the outcomes of their peers. Additional treatments relax anonymity and switch other-regarding concerns back on by implementing the linking preferences of a single "employed" participant.

We find broad support for strategic network formation. First, in treatments where other-regarding concerns are absent, agents are more likely to form new ties with peripheral peers, both when players' identities are known and when they are not. Our results are stronger when we focus the analysis on individuals who have performed well in an initial understanding test. Results from an additional treatment rule out priming effects due to the understanding test and the extensive explanation of incentives. Second, in treatments where other regarding considerations are made salient, we are unable to find evidence for our hypothesis. The answers to a post-play questionnaire suggest that many participants choose their links according to a norm of horizontal equality which does not reflect different endowments of network connections. Finally, in non anonymous treatments, agents link with peers whom they known in real life, even when this brings no additional material benefit to either party.

Overall, subjects seem able to understand the incentives that arise from the structure of the network and the referral process and are willing to adjust their linking decisions to maximise their chances of getting a referral. This central finding can be the basis for the design of hiring policies which strengthen the position of peripheral individuals in job contact networks.

This study is related to the empirical literature on social interactions in the labour market. This literature has deployed different empirical strategies to document significant peer effects in labour market outcomes in both developed and developing economies (Topa, 2001; Bayer et al., 2008; Magruder, 2010; Cingano and Rosolia, 2012), has identified important non-linearities in these effects (Beaman, 2012), has highlighted how referees respond to pecuniary incentives and may sometimes act opportunistically (Fernandez and Castillas, 2001; Fafchamps and Moradi, 2009; Beaman and Magruder, 2012) and how specific groups can be discriminated in the referral process (Beaman et al., 2013).

Our work also relates to the literature on network formation, which has so far presented some experimental evidence for strategic link formation (Callander and Plott, 2005; Conte et al., 2009), reflected on the role of inequity aversion (Goeree et al., 2009; Falk and Kosfeld, 2012), but also presented evidence consistent with a distate for equality (Van Dolder and Buskens, 2009).

Results from treatments where other-regarding considerations are made salient will be of interest for scholars studying the experimental evidence on other-regarding preferences and norms in Sub-Saharan Africa.(Barr and Stein, 2008; Jakiela, 2011; Miller Moya et al., 2011; Voors et al., 2011; Mueller, 2012).

Our work contributes to the literature in a number of ways. First, we provide the first test of the behavioural assumptions in the influential model of Calvo-Armengol (2004). Calvo Armengol's insight about the endogenous nature of social interactions in the labour market has been influential for the subsequent literature (Calvo-Armengol and Jackson, 2004; Wahba and Zenou, 2005; Galeotti and Merlino, 2010; Schmutte, 2012) and can prove very valuable in the design of policies to tackle economic exclusion. We show that the assumption of strategic network formation in the labour market passes a first, internally valid test. Second, our results further suggest that models of other-regarding preferences with high predictive power in simple allocation decisions can perform poorly in different domains- in our case, a link formation task. This is consistent with the evidence presented in Voors et al. (2011) and should motivate researchers to pay more attention to social norms that arise in specific domains. On the methodology side, we address typical endogeneity concerns that arise in dyadic settings through random link assignment. Furthermore, our design excludes non-equilibrium reasoning (Crawford et al., 2013), which may have shaped play in the repeated link formation games previously attempted in the literature, and hence produces clean evidence on self and other regarding motives in experimental link formation.

The next section presents the experimental design. Section 3 puts forward a number of predictions from a model of network formation with self and other regarding concerns. Section 4 presents the data and section 5 the results. Section 6 concludes.

2 Design

In the experiment, subjects are assigned to an exogenous "lab-network" and specify two additional peers they would like to be linked to. After subjects express their preferences, "lab-jobs" are assigned through a lottery. Each job-holders has to perform a task in the lab, for which he or she will be remunerated. Furthermore, each job-holder has to refer one unemployed contact in the lab-network for the job. If a job-holder has more than one unemployed peer, the referral is given to one of the eligible peers chosen at random. The network thus determines who can refer whom for a lab job.

In each lab session nine subjects play the game. Participants typically reside in the same neighborhood, often on the same block. Information about individuals real life connections is also exploited in the analysis.

The game proceeds as follows. First, each subject draws an ID letter from a urn. This letter remains private throughout the game. Second, each subject plays a standard dictator game with an anonymous opponent in the room. Third, instructions for the second part of the game are given, and subjects' positions in a network of undirected links, which is represented in figure 1 below, are revealed to them. Nodes are identified with letters and network ties are called "lab-friends". Subjects' understanding of the network structure and the incentives of the game is tested by means of a simple questionnaire. If more than one participant makes more than one mistake, the lab assistant is instructed to go through the explanation one more time. After understanding has been ensured, all subjects are asked to specify two additional agents to whom they would like to link.[5] Our analysis focuses on these linking decisions. Jobs are then drawn by the lab assistant. The network is

[5]If they do not specify a peer, or if the write R, a peer is randomly drawn for them.

updated according to a rule which varies with treatment. Referrals are assigned according to the updated network. While the lab assistant performs these tasks, participants are invited to respond to a short questionnaire on the motivations behind their choices in the experiment.

At the end of the game, job-holders are asked to perform the lab task and can then collect their winnings. These will include a show up fee, allocations from the dictator game, and the wage for the lab-job.[6]. Participants who got referred for a lab-job collect the show up fee and the dictator game allocation, and are then invited to come back the next day in order to perform the lab job and be paid for it. Finally, participants who did not get a job nor a referral collect their winnings and leave. All payments are given privately to the participants.

Treatments vary the network updating rule in order to identify different motives behind linking choices. In a first set of treatments, which we call SELF treatments, we update the network with the links specified by a single, randomly drawn unemployed player. This rule achieves two things. First, the additional links do not affect who the agent can give a referral to. This minimizes other regarding considerations.[7] Second, the fact that we implement the linking preference of a single player removes strategic thinking (Crawford et al., 2013). Subjects do not have to speculate about what others will do: if their choice is implemented, it will be the only modification to the existing network.[8]

[6]The show up fee was 0.5 USD, the total amount to be divided in the dictator game was about 1.1 USD and the wage 2.2 USD.

[7]While other regarding reasons are minimized, they are not wholly removed for sophisticated agents who care about the welfare of the other links of their chosen new partner. A new link to j decreases the chances that j's current friends are going to get a referral. While all links will impose a negative externality on two step away agents, an intrinsically motivated- and quite sophisticated!- agent may prefer to impose such externality on the better-off agents in the experiment. In the network we impose, this gives again a reason to link with degree one agents, as their only tie is a "well-off" degree 3 agent. While we cannot perfectly control for this type of reasoning, we offer two pieces of evidence which suggest that other regarding reasons are not influencing play in the SELF treatments. First, when given the opportunity to directly benefit the least well off partners in a different treatment, agents do no show a systematic desire to do so. It is not very plausible then that other-regarding considerations towards one-step away partners do not drive behaviour, while other-regarding considerations towards two-step-away partners do. Second, we interact linking decisions in this baseline treatment with the amount sent in the dictator game and show that there is no significant effect.

[8]Level 1 rationality players, for example, will worry that other subjects are also going to choose to link with degree one agents and that hence degree one agents will actually be quite central in the final network. Level 2 rationality players are going to best respond to level 1 players, and so on. Our procedure

In a second set of treatments, the OTHER treatments, we update the network with the links of a single, randomly drawn employed player. Linking decisions determine who will get the player's referrals and cannot be used to maximize the chance of getting a referral for oneself. Hence other-regarding motives become salient, while self-regarding considerations are switched off.

Treatment SELF and OTHER are played both with anonymous identities (SELFa, OTHERa) and with identities that are common knowledge (SELFn, OTHERn). In the latter, players are first asked to communicate their name in front of the group. Names are then written next to the respective node in the network and each participant is given a copy of this network map. Non-anonymous treatments give a more arduous test to the hypothesis that individuals prefer to link with peripheral peers. While centrality is clearly salient in the anonymous treatments, agents may focus on a number of other peer characteristics in the non-anonymous case.[9] Notice however that decisions remain private: at the end of the game participants are told whether they received a referral or not, but are not informed about how the network has been updated. This makes it hard for agents to require side-payments from each other after the experiment has been played.

In the standard protocol subjects are explicitly informed about the incentives that arise from the structure of the game.[10] Before asking for their linking decisions, players' understanding of the network map and of the incentives is tested with 5 questions.[11] If more than one participant makes more than one mistake, the lab assistant is instructed to go through the explanation one more time. Such explanations and tests are important to ensure participants understand the consequences of the decisions they are making. However, we worry that if participants have a desire to please the experimenter (Levitt and List, 2007; Zizzo, 2010), they may be primed by our questions to behave in the way they think we want them to behave. Our last treatment (SELFa2) is thus devised in order to reduce priming effects. In SELFa2, we give participants the same explanation of the rules of the game, but

rules out these considerations.

[9]For example, they may choose links which reinforce previous bonds with people they know, or they may choose to avoid individuals of specific social categories.

[10]Subjects in SELF treatments are told that a low degree agent is more likely to give them a referral than a high degree agent. Subjects in the OTHER treatments are told that a low degree agent is both the least likely player to get a referral for himself, and is the one whose chance of getting a referral increases the most with an additional link.

[11]These are presented in the appendix

Table 1: Summary of treatments

	Salient motive	Anonymity	Control for priming
SELFa	self-regarding	✔	
SELFa2	self-regarding	✔	✔
SELFn	self-regarding		
OTHERa	other-regarding	✔	
OTHERn	other-regarding		

omit any discussion of the incentives that these rules produces. Furthermore, we post-pone the test of understanding until after the players have made their linking decisions. [12] In this way we reduce both cognitive and social-related priming (Zizzo, 2010). Furthermore, the SELFa2 treatment allows us to test to what extent individuals in our sample independently grasp the incentives that derive from the asymmetric structure of the network and rival referral opportunities. Table 1 below summarizes the treatments and their characteristics.

Subjects are presented the network structure through a network map, which is reproduced in figure 1 below. Recent research in network cognition has uncovered a projection bias which is relevant to the present investigation: subjects overestimate the degree of agents characterised by a degree lower than themselves (Dessi et al., 2012). If that is the case among our experimental subjects, the role of centrality may become less salient, simply because high degree subjects underestimate the low degree of others. We take steps to limit this problem. First, the simple network structure in our design mitigates this concern. Second, the degree of each node is specified next to the ID letter. Third, the first three pre-play questions test whether subjects could infer network centrality from the map. Incorrect responses to these questions are extremely infrequent.

At the beginning of the experiment subjects play a standard dictator game (Camerer, 2003). Players split 20 Ethiopian Birr with an anonymous partner in the same session. Each subject is assigned to two different pairs: he plays the sender in first pair case and the receiver in the second pair. Players are aware that they are

[12]That is, we do not explain to participants that in our game a high degree agent is less likely to provide them with a referral. Agents can of course infer this from the rules of the game, and many of them do.

not playing twice with the same partner. Allocation decisions are private. Subjects are informed of the amount they have received only at the end of the lab session.

Dictator games offer a simple way to measure social preferences (Camerer, 2003). These are typically categorized as four standard groups (Charness and Rabin, 2002): selfish, competitive, inequality averse, and social welfare maximizing. Selfish individuals care only about maximizing their own payoff. Competitive individuals maximize the difference between their payoff and that of the receivers. Inequality-averse individuals minimize payoff differences between themselves and the other players. An inequality-averse person hence gives to others only when he or she starts with a better endowment. Finally, social welfare maximizers care positively about the receivers' payoffs regardless of the relative size of initial allocations.

In our experiment the dictator is endowed with ETB 20, whereas the receiver has no initial endowment. The amount given in this game hence reflects the strength of the sender‚Äôs concern for the receiver‚Äôs payoff when the initial allocation favors the sender. Giving in the dictator game is hence consistent with a preference for social welfare maximization and with inequity aversion. For the purposes of our experiment, we do not need to distinguish between these two types of preferences, as they would motivate identical behavior in the link-formation game.[13] We hence consider giving in the dictator game as a simple measure of pro-social social preferences.

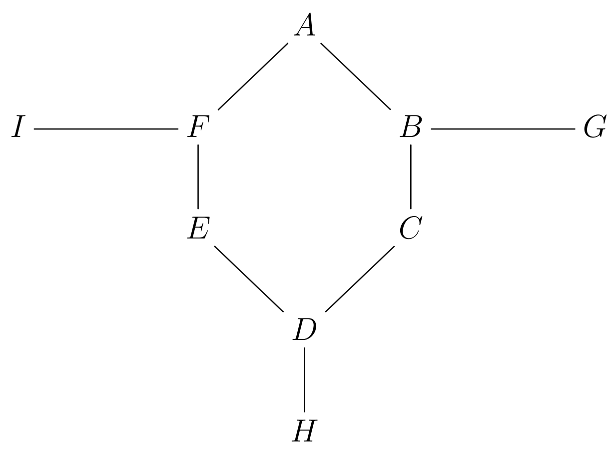
To ensure comparability and minimize noise factors during play, we follow a number of establishes practices in the lab-in-the-field literature. These include extensive piloting, simple standardized instructions that are read out to participants, double translation of all written material, and reliance on physical randomization devices (Barr and Genicot, 2008; Viceisza, 2012).

2.1 The Network

Each of the 9 participants in a session is assigned a position in the network presented in figure 1 below. In the network, there are three degree 1 agents, three degree 2 agents and three degree 3 agents. No agent has links with all agents of a given degree. Hence, when considering which additional links to establish, each agent in the network has at least one option for each centrality category.

[13]With one exception, explored in note 21.

Figure 1: The experimental network



3 Hypotheses

We now discuss in more detail the main hypotheses that will be tested.

Our first hypothesis is inspired by the model Calvo-Armengol (2004), whose strategic environment is reproduced in our game. Let $\pi_i(g)$ be i's chance of being employed given network g, $p \in (0,1)$ be the likelihood of being initially selected to perform the lab-job and $Q_i(g)$ be i's chance of being referred into the job by at least one peer in network g. Furthermore, let $N_i(g)$ be the set of direct contacts of i in g, and $n_i(g)$ be the number of contacts. $N_i^c(g)$ represents the set of missing direct links and $g+ik$ is the original network augmented by link ik. The strategic network formation model posits that individuals will choose link ik so as to maximise their chance of getting a lab-job. Formally, participants solve the following problem:

$$\max_{ik \in N_i^c(g)} \pi_i(g+ik) = p + (1-p)Q_i(g+ik) \tag{1}$$

$Q_i(g+ik)$ can be expressed as the inverse of the probability of receiving no referrals:

$$Q_i(g+ik) = 1 - \prod_{j \in N_i(g+ik)} (1 - pq(n_j(g)) \tag{2}$$

Conditional on k being employed, $q(n_k(g))$ gives the probability that k will refer i out of his $n_k(g)$ friends. This probability can be expressed as:

$$q(n_k(g)) = \sum_{z=0}^{n_k(g)-1} \binom{n_k(g)-1}{z} \frac{p^{n_k(g)-1-z}(1-p)^z}{z+1} \tag{3}$$

$$= \frac{1 - p^{n_k(g)}}{(1-p)n_k(g)} \tag{4}$$

After some algebra, (4) follows from (3).[14] When k is employed, the probability he will refer i boils down to the probability that i will be picked out of the average number of unemployed neighbors of k- $\frac{1}{(1-p)n_k(g)}$- conditional on at least one person being unemployed- $1 - p^{n_k(g)}$. It is easy to show algebraically that (4) decreases in $n_k(g)$:[15]

Proposition 1. $\frac{\partial q(n_k(g))}{\partial n_k(g)} < 0$ *for* $p \in (0, 1)$ *and* $n_k(g) \geq 0$

The result is intuitive: two links away partners are competitors for rival referrals opportunities. A partner with higher degree centrality is hence less likely to provide i with a referral. A subject who is solving problem (1) will thus prefer to link with the least central individual in g^c. Furthermore, in our game gains from following the optimal strategy are non-trivial. Conditional on being un-employed, a degree one agent who establishes two links with two other degree one players has a 72 percent probability of being drawn for a referral. This probability drops to 58 percent if he chooses to link with two degree 3 agents. This is a substantial 14 percentage points difference.[16] We thus formulate our first hypothesis:

Hypothesis 1. *Subjects in SELF treatments will choose new links with degree 1 peers*

Notice that this hypothesis is not consistent with models where network formation does not respond to extrinsic incentives. If subjects follow the simple heuristic of preferential attachment, for example, they will keep referring high degree subjects even if it is not in their material interest to do so.

Let use now turn to OTHER treatments. Now new links affect the employment chances of others, and do not affect the employment chances of ego. In our setting, the chances of employment increase monotonically with the number of direct links.

[14]See appendix for all proofs.

[15]This result is essentially the first part of Remark 1 in Calvo-Armengol (2004).

[16]Subjects in SELFa and SELF treatments are informed of these probabilities. Subjects in SELFa2 are not.

Degree 1 agents are hence those who have the smallest chance of employment. Furthermore, the marginal benefit of a new link is decreasing in the number of existing links. Degree 1 agents are hence also those who stand to gain the most from an additional link. More formally:[17]

Proposition 2. $\frac{\partial Q_i(n_k(g))}{\partial n_i(g)} \geq 0$ *and* $\frac{\partial^2 Q_i(n_k(g))}{\partial n_i(g)^2} \leq 0$

There is now an established literature in economics which explores other-regarding preferences. In a frequently cited paper, Charness and Rabin (2002) categorize the standard types.[18] Players who are social welfare maximizers and inequity averse would send money in a dictator game which starts with unequal endowments.

For illustration, let us assume other-regarding preferences of the following form:[19]

$$u_i(g) = \pi_i(g) + \frac{\gamma}{n-1} \sum_{j \in N} \pi_j(g) \tag{5}$$

γ is the altruism parameter. Any ik link will increase i's and k's welfare, while decreasing the welfare of the $n_i(g)$ current contacts of i, who are now facing one more competitor for i's referral. We can hence decompose the effect of the linking decision on i's utility in three elements:

$$\begin{aligned} u_i(g+ik) - u_i(g) =& \pi_i(g+ik) - \pi_i(g) \\ &+ \frac{\gamma}{n-1} \pi_k(g+ik) - \pi_k(g) \\ &+ \frac{\gamma}{n-1} \sum_{j \in N_i(g)} \pi_j(g+ik) - \pi_j(g) \end{aligned} \tag{6}$$

The first element, reflecting the payoff of i in the game, is not affected by which particular k is chosen. The third element reflects the negative externality on i's current links. Such externality is present no matter whom i decides to link to and

[17] The second part of Remark 2 in Calvo-Armengol (2004) makes a point similar to the first part of proposition 2 here.

[18] Inequity averse agents feel envy towards those with higher payoff and guilt towards those with lower payoff. Social welfare maximizers care positively about the welfare of the other person, without conditioning on relative payoffs. Competitive agents maximise the difference between oneself and the other players. Selfish agents maximise one's own payoff.

[19] These are akin to the social welfare maximising type.

does not depend on $n_k(g)$.[20] The second term is what motivates the agent to link with the less central individuals. Proposition 2 shows that the marginal payoff k gets from a link with i decreases in k's degree. Choosing the least central k means establishing the link with the agent that will benefit from the connection the most. This maximises the value of the second term in (6). Hence we hypothesize that subjects in the OTHER treatment will be more likely to link with their peripheral peers.[21]

Participants may have an additional reason to include peripheral peers¿ the fact they are the least well off in the game. Several questionnaire studies in empirical social choice have reported that other-regarding agents attach special weight to the well-being of the least well off (Yaari and Bar-Hillel, 1984; Gaertner and Schokkaert, 2011). These considerations would strengthen the desire to link with degree 1 individuals.

Hypothesis 2. *Subjects in OTHERa and OTHERn treatments will create new links with degree 1 peers*

This effect is driven by pro-social, other regarding individuals. Hence we expect a positive correlation between giving in the dictator game, which is meant to capture such preferences, and the choice to link with a degree 1 peer. Notice that while giving in the dictator game is costly, the choice to include a peripheral peer does not involve any material sacrifice. This minimizes concerns about wealth effects generating a spurious negative correlation between the two games (Andreoni and Miller, 2002).

Hypothesis 3. *Subjects who have sent a positive amount in the dictator game are more likely to link with a degree 1 peer in the OTHERa and OTHERn treatments. Dictator game giving is uncorrelated with play in SELF treatments.*

[20] Notice that if the player specifies no link, two links are assigned to him anyways. Also notice that unemployed agents can receive multiple referrals, in which case one of the referrals is simply lost. This ensures that the degree of k has no influence on the size of the negative effect on the current contacts of i.

[21] A similar argument can be made for inequity averse agents. In the case of inequity aversion, however, there is an ambiguity for the behaviour of degree 1 players, who may fear becoming envious of their degree 1 peers if they link with them. The empirical analysis reveals degree 1 players behave no differently in the game and we hence explore this point no further.

The game is played by subjects who come from the same neighborhoods and, in many instances, from the same block. Theories of directed altruism predict that individuals in the OTHERn treatment act more altruistically towards acquaintances (Goeree et al., 2010; Leider et al., 2009; Ligon and Schechter, 2012). We do expect this effect to be at work in the OTHERn treatment, where agents are aware of the identities of those who receive the referral. We do not expect, however, directed altruism to play a role in the choices under the SELFn treatment, where links to a friend do not produce any material benefit for the friend and are sometimes detrimental to one's own material benefit.[22]

Hypothesis 4. *Subjects in the OTHERn treatment will be ceteris paribus more likely to refer those whom they know in real life. Decisions of subjects in the SELFn treatments will not be affected by such knowledge.*

We will analyze the data using dyadic regression analysis. In particular, we will test hypotheses 1 and 2 using models of the following form:

$$r_{ij} = \alpha + \beta_1 c2_j + \beta_2 c3_j + e_{ij} \tag{7}$$

The unit of observation is all initially unlinked, directed dyads.[23] r_{ij} is a dummy which takes value 1 if i has chosen to establish a new link with j. cx_j is another dummy variable indicating individuals with experimental degree centrality x. The coefficients on cx_j will provide the basic test for hypotheses 1 and 2. If β_1 and β_2 are significant and negative we will have evidence that in our experimental social network agents value peripheral partners more than central ones.

We run separate analyses for the data from the SELF and OTHER treatments, as these elicit different decision mechanisms. We investigate the effects of sub-treatments and of i's characteristics such as understanding or giving in the dictator game by means of simple dummies interacted with j's degree:

$$r_{ij} = \alpha + \beta c2_j + \gamma c3_j + \delta t_i + \theta(t_i * c2_j) + \lambda(t_i * c3_j) + u_{ij} \tag{8}$$

[22] This is the case when the friend has degree 2 or 3.

[23] This means that the matrix is not the full $n(n-1)$ square matrix. We call the dyads directed because they express directed willingness to link. Notice that the actual dyads in the network are undirected

Models (7) and (8) will be estimated using OLS, correcting standard errors for arbitrary correlation at the session level. Previous studies have shown that when the number of independent groups of observations is low, which is often defined as less than 42, hypothesis tests based on clustered standard errors over-reject the null. Our regression analysis is based on 30 sessions of the SELF treatment and 20 sessions of the OTHER treatment and is hence characterized by a low number of clusters. In a widely cited paper Cameron et al. (2008) show that the wild bootstrap-t method can be used to achieve accurate inference even with few clusters. This method simulates a distribution of the test statistics which can be used for hypothesis testing in conjunction with the original test statistics. We apply the wild-bootstrap-t method throughout the analysis. As reccomended by Cameron et al. (2008), we impose the null hypothesis of no effect and use Rademacher weights for resampling. Regression tables will report p-values obtained from the bootstrapped distribution of the test statistics.

4 Data

The fieldwork for this project took place between September and October 2012 in the city of Kombolcha, on the main road between the capital Addis Ababa and Mekelle, in the South Wollo province of Amhara region. According to the 2007 census, the city has a population of about 59,000. In recent years, Kombolcha has benefited from a number of industrial developments. The city can now count on an expanded textile factory, a metalwork factory, a large brewery, as well as smaller firms working on the processing of leather and seeds. This expansion of the formal sector makes Kombolcha an ideal place to study job referral networks in Ethiopia. As constraints on the number of available jobs are progressively relaxed, it is important to investigate whether some individuals and groups are excluded from the allocation of the new economic opportunities. Background qualitative fieldwork at the onset of the project and descriptives from the survey data reveal that reliance on job contact networks in Kombolcha is indeed extensive.

Our sampling strategy was based on the following steps. First, based on qualitative fieldwork and discussions with local officials, we identified three low-to-middle income urban, residential neighbourhoods.[24] We delineated all residential blocks of houses in each neighbourhood using the Google-Earth map of the city and randomly sampled 19 blocks. In each block, we listed all individuals in the age group

[24]We excluded the other three neighbourhoods either because they included rural agglomerates or because they had few dwellers.

20-40 resident in the block and in town at the time of our fieldwork. We invited all sampled individuals to take part in the experiment. 447 individuals took part in the experiment, out of 518 that were interviewed.[25]

The sampling strategy enables us to capture block level networks. We have reason to believe this is a focal domain of interaction. There is substantial support for this assumption in the empirical literature.[26] Furthermore, in our sample, we measure an exchange of information or referrals for a non-trivial 21 percent of dyads.[27]

Network information was collected in the following way. Individuals were presented with a list of all people in the 20-40 age group residing in the block and were asked to identify those people they knew. This corresponds to the star induced network defined by Chandrasekhar and Lewis (2012).[28] For each link, respondents were asked questions regarding the strength of the link and various dimensions of social interaction: asking, giving and receiving job information; giving and receiving referrals; borrowing and lending; gift exchanges. Subjects that knew each other had spoken on average on 12 days of the previous moth. In about 65 percent of

[25]Table 4 in the appendix shows that there are few statistically significant differences in observable characteristics between individuals who took part in the experiment and individuals who did not. Selected individuals are less likely to be Muslim and tend to report a higher number of links in the neighborhood. When we replace i's self-reported links with the number of peers who have mentioned i as a friend, the selection effect on the network variable disappears.

[26] Marmaros and Sacerdote (2006), for example, show that geographical proximity is an important determinant of friendship among US adolescents. Their empirical strategy is particularly credible as it relies on random assignment to dorms. In a developing country setting, Karlan et al. (2009) report that 59 percent of observed dyadic ties among Peruvian shanty town dwellers are between neighbours, while Fafchamps and Gubert (2007) document that geographic proximity is a strong predictor of risk sharing behavior. Looking more specifically at labour market settings, Bayer et al. (2008) analyse census data on the Boston metropolitan population and find a strong, significant effect of shared block residence on the probability of working in the same census tract. Similarly, Hedstrom et al. (2003) document peer effects in the duration of unemployment among Stockholm youth. Their proxy of peer group is also given by small geographical units where a median number of 66 young people reside. Finally, Topa (2001) shows that unemployment pattern in Chicago's neighbourhoods are consistent with a peer effect model, albeit he focuses the analysis on geographical areas far larger than blocks.

[27]Overall, our empirical dataset consists of 15, 588 block-level dyads among 518 individuals in 19 blocks. In 1,804 cases i knows j and in 377 cases i declares to have given to or received from j job information or a referral.

[28]We included in the list individuals who reside in the block but were not available for interview (in most cases, they were out of town) at the time of fieldwork

cases, i defined j as a "worship place acquaintance", in 12 percent of cases as a member of the same family and in 14 percent of cases as a close friend. Figures 4, and 5 in the appendix present this descriptive data.

In figures 2 and 3 we briefly summarise data about job contact networks. In this figures, an undirected job contact link is said to exist if i has reported any exchange of information or referrals from i to j, or from j to i. Figure 2 shows the architecture of one of the 19 blocks in our sample. Blue squares represent individuals. Lines represent the job contact links. The isolated squares on the left of the picture represent the individuals with no block-level job contact links. We define the number of job contact links an individual has as his or her network degree. In figure 3 we plot the distribution of the network degree of the respondents.

The survey also included standard socio-demographic such as household characteristics, age, gender, ethnicity, education, and migration status. A detailed module on labour market experience was administered, capturing employment status, job characteristics, search strategies (while in unemployment and on-the-job search) and referrals. A further module investigated expectations regarding employment, wage and unemployment exit rates.

We report here some summary statistics for the characteristics of our experimental sample.[29] We are not able to match 16 IDs recorded in the experimental forms with a questionnaire.[30] We hence have 447 observations for experimental variables, and 430 observations for individual variables.[31]

We check for covariates balance for both levels of randomization. Tables 5 and 6 in the appendix show the result of our tests employing simple OLS regressions. Each column corresponds to a regression of a different dependent variable on dummies

[29] We categorize workers as employed, unemployed, and inactive. A worker who currently has a job is classified as employed. A worker who does not have a job, has been searching for one in the past seven days, and is currently available for work is defined as unemployed. A worker who does not have a job and either is not available for work or has not been searching for work in the past seven days (or both) is considered inactive. Also we report a variable called "block network degree". This is defined as the self-reported number of social ties with residents of the same block in the age group 20 to 40.

[30] While we have no particular reason to suspect that these individuals infringed the rules of the game in any way, we still exclude them from the analysis presented below. Inclusion of these 16 observations does not change any of the results.

[31] Furthermore, 1 individual declined to respond the question about religion and 9 individuals declined to respond the question about migration status.

Table 2: Summary statistics: Binary variables

Variable	Proportion	Std. Dev.	N
Male	0.488	0.5	430
Muslim	0.536	0.499	429
Migrant	0.223	0.417	421
Employed	0.388	0.488	430
Inactive	0.374	0.485	430

Table 3: Summary statistics: Continuous variables

Variable	Mean	Std. Dev.	N
Age	26.926	9.404	430
Education	9.606	4.394	430
Earnings	433.356	801.825	430
Block network degree	6.047	4.088	430
Amount kept in DG	13.938	4.158	430

for individual and session treatments. The column heading specifies the dependent variable that is being tested. Table 5 confirms that the observable characteristics of participants assigned to different levels of network centrality are not statistical different. Table 6 shows that there are some weak differences across individuals in the SELF and OTHER treatments: individuals in the SELF treatments are more likely to be male, older and are less central in their real world network. These affect are not large and only significant at the 10percent level, however, they motivate caution when comparing the magnitudes of effects across treatments. In the analysis that follows, we include controls for these characteristics.

5 Results

Result 1. *Understanding is high and uncorrelated with session level treatment. Understanding in the SELFa2 session is no lower than in the SELFa and SELFn sessions*

We tested understanding with 5 questions. For these we have complete responses for 444 subjects out of 447. The first three questions dealt with understanding of the network graph. Very few people got these questions wrong. The further two questions tested understanding of the relevant incentives.[32] There was somewhat more variation here and we hence create a binary variable for whether the participant answered both of these questions correctly. Reassuringly, about 80 percent of participants chose the right answer in both questions.

Table 7 in the appendix shows the result of a linear probability model where the understanding variable is regressed over a number of session treatment dummies.[33] The first column includes dummies for four treatments: SELFa is the residual category. There are no significant differences in understanding across treatments. The second column shows that in the SELFa2 treatment, where no explanation of the relevant incentives was given, understanding of the incentives resulting from the structure of the network was not significantly lower than in other treatments. 77 percent of participants in the SELFa2 treatment answered both questions correctly.

If participants did not want to make a linking decision, they had the option of writing the letter R, in which case a link would be picked at random for them.[34] A high percentage of random decisions could be interpreted as a signal of poor understanding. The data dispels such concern. The random link option was used in only 12 percent of decisions and only 8 percent of participants chose a random link twice. Furthermore, the likelihood of choosing a random link was not statistically different across treatment.

Taken together, these results reassure us that the experiment was well understood. We can hence proceed to analyse the linking decisions of participants. Table 8 in the appendix gives frequencies and percentages of the linking decisions made in the experiment.

Result 2. *Subjects in SELF treatments are more likely to establish links with peripheral individuals than with more central peers*

[32] In the SELF treatments, this was the probability of receiving a referral from agents with different degree centrality. Participants were asked whether they would be more likely to receive a referral from a degree X or a degree Y agent. In the OTHER treatments we focused on the probability that an agent with a given level of degree centrality would get a referral for himself. We asked participants whether an agent of degree X was more likely to get a referral than an agent of degree Y.

[33] Standard errors are clustered at the session level, as in all other regressions in this paper

[34] Participants could write R in place of either of the links. A blank box counted as a letter R

Table 9 below shows results from estimation of the dyadic regression model 7. Facing a degree 2 or degree 3 player significantly decreases the probability of referral, compared to the residual category of degree 1 players. Quantitatively, a connection to a degree 2 partner is 17 percentage points less likely, and a connection to a degree 3 partner is 20 percentage points less likely, than a connection to a degree 1 partner. These results are a broad confirmation of hypothesis 1. Participants understand the incentives arising from competition for referrals in a network where links are distributed unequally and their linking decisions are consistent with a concern to maximise the chance of referral. A simple descriptive analysis of the post-play questionnaire is consistent with our interpretation of the results: 75 percent of players in SELF treatments answer that they played to maximise the chance of getting a referral. Figure 7 in the appendix shows this data graphically. However, not all participants choose to link with peripheral agents in the SELF treatments. This may be consistent with heterogeneity of motives or mistakes.[35]

< Table 9 here >

In column 3 we test whether play differs when anonymity is relaxed or incentives are not explained. The latter has no statistically significant effect on the likelihood of linking to a peripheral individual. This is evidence against priming, and is consistent with responses to the post-play questionnaire: individuals in SELFa2 were even more likely to report that they played to maximise their chance of referral than their counterparts in SELFa. Furthermore, this shows that individuals in our sample are able to grasp the strategic incentives which result from competition for referrals in a network with agents endowed with different levels of centrality. Second, relaxing anonymity increases somewhat the number of degree 3 individuals who are chosen. This is possibly due to additional motives for linking in non-anonymous treatments: about 7.5 percent of players in SELFn claim that they have chosen their real friends.[36]

Result 3. *The likelihood to link with a peripheral agent in SELF treatments is significantly higher for players with high understanding of the game. It is not correlated with giving in the dictator game.*

Result 3 focuses on heterogeneity of play. In table 10 below we show that main effect is unchanged when restricting attention to high understanding players

[35]We plan to do more work on this area.

[36]We further explore this effect in the analysis below.

(column 1).[37] We further show that high understanding players are less likely to choose degree 2 and degree 3 partners than low understanding players (column 2), and that participants who sent a positive amount in the dictator game, which is consistent with inequity aversion or social welfare maximization, do not respond differently to the centrality of potential partners (column 3).

< Table 10 here >

Let us now turn to subjects in OTHER treatments. We first look at the amount sent in the dictator game. Figure 6 summarises the data: giving is substantial. The mean amount sent is 6 Ethiopian Birr, which is just a third of the endowment, and is in line with the experimental evidence across the world (Camerer, 2003). We interpret this result as evidence of substantial other-regarding motives in the population. Furthermore,we note that the modal amount sent is half of the endowment. This potentially reflects a a strong preference for or norm of equality.

< Figure 6 here >

Result 4. *The mean amount sent in the dictator game is a third of the endowment. The equal split of the endowment is the modal choice.*

We now estimate model 7 over data from the OTHERa and OTHERn conditions and report the results in table 11. The coefficients on the dummies for j's degree are small and insignificant. Subjects in non anonymous treatments seem less likely to link with peripheral peers than subjects in anonymous treatments, but this difference is not significant. If we restrict the sample to subjects in the OTHERa treatment, the effect has the hypothesized direction, but is statistically imprecise. In short, we are unable not find evidence for hypothesis 2. Individuals do not seem to link with peripheral peers more often than if they were choosing with a random rule.

In table 12 we test whether we can find limited support for hypothesis 2 for individuals who understood the game well. This turns out not to be the case. We

[37] In this regression model we omit the dummies for centrality 2 and centrality 3 and focus on the effects of these two degrees of centrality only when interacted with high understanding. This amounts to considering an ij dyad where j has centrality 2 or 3 and i has low understanding as untreated.

also cannot find evidence that giving in the dictator game is correlated with the decision to link with a peripheral peer.[38] This falsifies hypothesis 3.

Result 5. *Subjects in OTHER treatments are NOT more likely to link with peripheral peers. Linking decisions are not correlated with the degree of the new tie.*

Result 6. *Linking decisions in OTHER treatments are uncorrelated with understanding or giving in the dictator game*

< **Table 11 here** >

< **Table 12 here** >

Responses from the question about players' motivation are helpful to explain the above results. In OTHERa, where reasons related to the network position should be more prominent, about 29 percent of individuals declared that they "chose at random" in order to give every peer an equal chance, while only 12 percent of individuals claimed to have "chosen the peer who needed help the most".[39] In other words, reducing inequality in links endowments was not perceived as the fair thing to do. This result is striking given that in allocation tasks subjects in our sample and in other studies across sub-Saharan Africa are typically willing to transfer own resources in a way that reduces or cancel initial inequalities in monetary endowments (Miller Moya et al., 2011; Mueller, 2012).

Finally, we analyze separately the data from the non anonymous treatments. Here we can use a full set of dyadic covariates. Our hypothesis is that knowing individual j will motivate individuals in the OTHERn treatment, but not in SELFn. Estimates are reported in table 13. Subjects in OTHERn are indeed more likely to link with individuals they know in real life. The effect is large, but is measured with little statistical precision: it is significant at the 10 percent level before bootstrapping and becomes insignificant after we apply the wild bootstrap-t method. Second, the result for SELFn is opposite to what was hypothesized. Linking decisions are

[38]Again, we are omitting the centrality 2 and 3 dummies. So we are considering all ij dyads where i has low understanding (column 1), has not sent anything in the dictator game (column 2) or both (column 3) as untreated. The coefficients on the interactions tell us whether a ij connection in treated dyads where j has centrality 2 and centrality 3, respectively, is less likely than a connection in untreated dyads.

[39]Figure 8 in the appendix shows self-reported motives for the OTHERa treatment.

still significantly driven by material concerns, as showed by the significant and large coefficients on the dummy variables for the degree of player j. However, individuals are also more likely to link with their actual peers. Personal knowledge of j raises the probability of a link by about 10 percentage points. This is surprising as in this game choosing somebody you know brings no benefit to yourself or to your peers. This result is instead consistent with a model where decision makers exhibit concerns for social identity and have a preference to re-establish such identity by linking to in-group members.

Result 7. *In the SELFn treatments, subjects are more likely to link with known peers*

< Table 13 here >

We also test whether in OTHERn individuals are more likely to link with known peers who earn less, do not have a job, or have fewer connections in the neighborhood. These effects would be consistent with the model of other regarding preferences outlined above and would indicate that individuals are responding to field characteristics and not to the characteristics imposed in the lab. We are unable to find any evidence of these effects.[40]

In the post-play questionnaire, only about 3 percent of players in SELFn and OTHERn answer that they chose people they knew in real life. However, participants may have felt embarrassed to admit making such discriminations. Furthermore, statistical power may be a problem in this last set of regressions on non-anonymous treatments: i knows j in only about 8.5 (7) percent of dyads in SELFn (OTHERn).

[40] Regression tables are not included for concision, but are available upon request.

6 Conclusion

In poor, growing economies like Ethiopia structural change and the increased availability of well-paying jobs hold a promise to deliver substantial poverty reduction in the coming years. However, individuals and groups that perform poorly in the search process are less likely to benefit from these new opportunities (African Development Bank, 2012; World Bank, 2013). Job contact networks play a crucial role in this respect in many developed and developing countries, including Ethiopia. Using unique dyadic data on social interactions in the labour market of a mid-size Ethiopian town we reveal a skewed distribution of links in neighborhood level job market networks. Many individuals do not take part in the exchange of valuable information and referrals. This is inconsistent with extant models of strategic network formation, which posit that competition for scarce referrals motivates self-interested individuals to link with so far poorly connected peers. It is also inconsistent with models of other-regarding preferences, such as social efficiency maximisation or inequity aversion. However, with observational data alone it is hard to establish whether these predictions fail because the models misrepresent individuals' decision making process or because they do not capture the relevant incentives in the field. In this paper we devise a lab-in-the-field experiment which tests whether individuals link with peripheral peers when it is in their material interest to do so, and when other regarding considerations are made salient.

In treatment where agents form additional links to receive referrals and where competition for referrals is salient, we find broad support for strategic network formation. Agents are more likely to form new ties with currently peripheral agents and such effect is robust to the relaxation of the anonymity condition. Furthermore, subjects understand the incentives deriving from competition for referrals even when these are not explicitly explained and do not seem to be primed by our explanations into behaving strategically. In treatments where other regarding considerations are made salient, we are unable to find evidence of a significant desire to include peripheral individuals. This is despite the fact that most people give substantial amounts in an initial dictator game, which is suggestive of widespread other-regarding preferences in the population. Evidence from a post-experiment questionnaire tentatively explains this finding as the result of a norm of equality which is applied despite unequal endowments of network links. Finally, in all non anonymous treatments, agents have a tendency to link with peers whom they known in real life, even when this brings no additional material benefit to either party.

Overall, our evidence is consistent with the theory of strategic network formation put forward in Calvo-Armengol (2004). This is suggestive of a potential, context-dependent mechanism of inclusion. However, the lack of an unconditional social preference to support peripheral individuals suggests that when local incentives do not favour the latter, unequal access to job contact networks will persist. Observational data from the city of Kombolcha indeed shows that integration in job contact networks varies widely between individuals. To revert this situation, policy can target the field incentives to which network formation responds. For example, caps could be set to the number of referrals employers can ask from employees of established groups. Albeit such policy may sometimes result in adverse selection (Beaman and Magruder, 2012; Beaman et al., 2013), it would have beneficial effects on two levels. First, it would increase the number of peripheral referees and hence, due to homophily, the number of individuals from peripheral groups referred into jobs. Second, as suggested by the results of this paper, it would make the rivalry of referrals more salient, and hence incentivize workers to include poorly connected individuals in their job networks, breaking the cycle of exclusion.

References

Andreoni, J. and J. Miller (2002, March). Giving according to garp: An experimental test of the consistency of preferences for altruism. *Econometrica 70*(2), 737–753.

Barr, A. and G. Genicot (2008, December). Risk sharing, commitment, and information: An experimental analysis. *Journal of the European Economic Association 6*(6), 1151–1185.

Barr, A. and M. Stein (2008). Status and egalitarianism in traditional communities: An analysis of funeral attendance in six zimbabwean villages. Technical report.

Bayer, P., S. L. Ross, and G. Topa (2008, December). Place of work and place of residence: Informal hiring networks and labor market outcomes. *Journal of Political Economy 116*(6), 1150–1196.

Beaman, L., N. Keleher, and J. Magruder (2013). Do job networks disdvantage women? evidence from a recruitment experiment in rural malawi. Technical report.

Beaman, L. and J. Magruder (2012, December). Who gets the job referral? evidence from a social networks experiment. *American Economic Review 102*(7), 3574–93.

Beaman, L. A. (2012). Social networks and the dynamics of labour market outcomes: Evidence from refugees resettled in the u.s. *Review of Economic Studies 79*(1), 128–161.

Bramoulle, Y. and R. Kranton (2007). Risk-sharing networks. *Journal of Economic Behavior & Organization 64*(3-4), 275–294.

Callander, S. and C. R. Plott (2005, August). Principles of network development and evolution: an experimental study. *Journal of Public Economics 89*(8), 1469–1495.

Calvo-Armengol, A. (2004). Job contact networks. *Journal of Economic Theory 115*, 191–206.

Calvo-Armengol, A. and M. Jackson (2004). The effects of social networks on employment and inequality. *American Economic Review 94*(3), 426–454.

Calvo-Armengol, A. and Y. Zenou (2005, May). Job matching, social network and word-of-mouth communication. *Journal of Urban Economics 57*(3), 500–522.

Camerer, C. F. (2003). *Behavioral Game Theory.* New York, NY: Russell Sage Foundation.

Cameron, A. C., J. B. Gelbach, and D. L. Miller (2008, August). Bootstrap-based improvements for inference with clustered errors. *The Review of Economics and Statistics 90*(3), 414–427.

Chandrasekhar, A. and R. Lewis (2012). Econometrics of sampled networks. *mimeo*.

Charness, G. and M. Rabin (2002, August). Understanding social preferences with simple tests. *The Quarterly Journal of Economics 117*(3), 817–869.

Cingano, F. and A. Rosolia (2012). People i know: Job search and social networks. *Journal of Labor Economics 30*(2), 291 – 332.

Conte, A., D. D. Cagno, and E. Sciubba (2009, June). Strategies in social network formation. Birkbeck Working Papers in Economics and Finance 0905, Birkbeck, Department of Economics, Mathematics & Statistics.

Crawford, V., M. Costa-Gomes, and N. Iriberri (2013, September). Structural models of nonequilibrium strategic thinking: Theory, evidence, and applications. *Journal of Economic Literature* (1), 5–62.

Dessi, R., E. Gallo, and S. Goyall (2012, June). Network cognition. mimeo.

Fafchamps, M. and F. Gubert (2007, July). The formation of risk sharing networks. *Journal of Development Economics 83*(2), 326–350.

Fafchamps, M. and A. Moradi (2009, August). Referral and job performance: Evidence from the ghana colonial army. CEPR Discussion Papers 7408, C.E.P.R. Discussion Papers.

Falk, A. and M. Kosfeld (2012). It's all about connections: Evidence on network formation. *Review of Network Economics 11*(3), 2.

Fehr, E. and K. M. Schmidt (1999). A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics 114*(3), pp. 817–868.

Fernandez, R. and E. Castillas (2001). *How much is that network worth? Social capital in employee referral networks*, Chapter 4. Aldine-deGruyter.

Gaertner, W. and E. Schokkaert (2011). *Empirical Social Choice.* Number 9781107013940 in Cambridge Books. Cambridge University Press.

Galeotti, A. and L. P. Merlino (2010, May). Endogenous job contact networks. ISER Working Paper Series 2010-14, Institute for Social and Economic Research.

Goeree, J. K., M. A. McConnell, T. Mitchell, T. Tromp, and L. Yariv (2010, February). The 1ḷaw of giving. *American Economic Journal: Microeconomics 2*(1), 183–203.

Goeree, J. K., A. Riedl, and A. Ule (2009, November). In search of stars: Network formation among heterogeneous agents. *Games and Economic Behavior 67*(2), 445–466.

Granovetter, M. (1995). *Getting a Job. A Study of Contacts and Careers* (2nd ed.). The University of Chicago Press.

Harrison, G. W. and J. A. List (2004, December). Field experiments. *Journal of Economic Literature 42*(4), 1009–1055.

Hedstrom, P., A. Kolm, and Y. Aberg (2003, November). Social interactions and unemployment. Working Paper Series 2003:15, IFAU - Institute for Labour Market Policy Evaluation.

Ioannides, Y. M. and L. D. Loury (2004, December). Job information networks, neighborhood effects, and inequality. *Journal of Economic Literature 42*(4), 1056–1093.

J-PAL (2013). J-pal youth initiative review paper. Technical report.

Jackson, M. O. (2008). *Social and Economic Networks.* Princeton, NJ, USA: Princeton University Press.

Jakiela, P. (2011). Social preferences and fairness norms as informal institutions: Experimental evidence. *American Economic Review 101*(3), 509–13.

Karlan, D., M. Mobius, T. Rosenblat, and A. Szeidl (2009, August). Trust and social collateral. *The Quarterly Journal of Economics 124*(3), 1307–1361.

Leider, S., M. M. Mobius, T. Rosenblat, and Q.-A. Do (2009, November). Directed altruism and enforced reciprocity in social networks. *The Quarterly Journal of Economics 124*(4), 1815–1851.

Levitt, S. and J. List (2007). What do laboratory experiments measuring social preferences reveal about the real world? *Journal of Economic Perspectives 21*(2), 153–174.

Ligon, E. and L. Schechter (2012). Motives for sharing in social networks. *Journal of Development Economics 99*(1), 13–26.

Magruder, J. R. (2010, January). Intergenerational networks, unemployment, and persistent inequality in south africa. *American Economic Journal: Applied Economics 2*(1), 62–85.

Mano, Y., T. Yamanops, A. Suzuky, and T. Matsumotol (2010, November). Local and personal networks in employment and the development of the labour markets: Evidence from the cut flowers industry in ethiopia. GRIPS Policy research Centre Discussion Paper 10-29, National Graduate Institute for Policy Studies.

Marmaros, D. and B. Sacerdote (2006, 02). How do friendships form? *The Quarterly Journal of Economics 121*(1), 79–119.

African Development Bank (2012). *African Economic Outlook 2012: Promoting Youth Employment.* Oecd Development Centre. OECD Publishing.

World Bank (2013). World development report 2013: Jobs.

Miller Moya, L. M., A. Barr, J. Burns, and I. Shaw (2011, September). Individual notions of distributive justice and relative economic status. DFAEII Working Papers 2011-03, University of the Basque Country - Department of Foundations of Economic Analysis II.

Mueller, A. (2012). In the public eye. norms of distributive justice and sharing behaviour under asymmetric information: Evidence from rural malawi. Technical report.

Schmutte, I. (2012). Endogenous job referral networks: theory and empirical applications.

Serneels, P. (2007, 02). The nature of unemployment among young men in urban ethiopia. *Review of Development Economics 11*(1), 170–186.

Topa, G. (2001, April). Social interactions, local spillovers and unemployment. *Review of Economic Studies 68*(2), 261–95.

Topa, G. (2011). Labor markets and referrals. Volume 1 of *Handbook of Social Economics*, pp. 1193 – 1221. North-Holland.

Van Dolder, D. and V. Buskens (2009, May). Social motives in network formation: An experimental study. *International Conference on Game Theory for Networks. Papers and Proceedings* (1), 593–602.

Viceisza, A. (2012). Treating the field as a lab: A basic guide to conducting economics. Food security in practice technical guide series 7, International Food Policy Research Institute (IFPRI).

Voors, M., E. Bulte, A. Kontoleon, J. A. List, and T. Turley (2011, May). Using artefactual field experiments to learn about the incentives for sustainable forest use in developing economies. *American Economic Review 101*(3), 329–33.

Wahba, J. and Y. Zenou (2005, December). Density, social networks and job search methods: Theory and application to egypt. *Journal of Development Economics 78*(2), 443–473.

Yaari, M. and M. Bar-Hillel (1984). On dividing justly. *Social Choice and Welfare* (1), 1–22.

Zizzo, D. (2010, March). Experimenter demand effects in economic experiments. *Experimental Economics 13*(1), 75–98.

7 Appendix

7.1 Derivations

7.1.1 From equation (3) to (4)

$$\sum_{z=0}^{n_k(g)-1} \binom{n_k(g)-1}{z} \frac{p^{n_k(g)-1-z}(1-p)^z}{z+1}$$

Define $m \equiv n_k(g) - 1$, so that:

$$\sum_{z=0}^{m} \binom{m}{z} \frac{(p^{m-z}(1-p)^z}{z+1}$$

Now pre-multiply by (1-p)(m+1):

$$(1-p)(m+1)\sum_{z=0}^{m} \binom{m}{z} \frac{p^{m-z}(1-p)^z}{z+1}$$

$$(1-p)(m+1)\sum_{z=0}^{m} \frac{m!}{z!(m-z)!}\frac{p^{m-z}(1-p)^z}{z+1}$$

$$(1-p)(m+1)\sum_{z=0}^{m} \frac{m!}{(z+1)!(m-z)!}p^{m-z}(1-p)^z$$

$$\sum_{z=0}^{m} \frac{m+1!}{(z+1)!(m-z)!}p^{m-z}(1-p)^{z+1}$$

$$\sum_{z=0}^{m} \frac{m+1!}{(z+1)!(m+1-(z+1))!}p^{m+1-(z+1)}(1-p)^{z+1}$$

Now define $s \equiv z + 1$:

$$\sum_{s=1}^{m+1} \frac{m+1!}{(s)!(m+1-(s))!}p^{m+1-s}(1-p)^{s}$$

Notice that the Binomial theorem implies that: .

$$\sum_{s=0}^{m+1} \frac{m+1!}{(s)!(m+1-(s))!}p^{m+1-s}(1-p)^{s} = 1$$

Thus:

$$\sum_{s=1}^{m+1} \frac{m+1!}{(s)!(m+1-(s))!}p^{m+1-s}(1-p)^{s} = 1 - \frac{m+1!}{0!(m+1-0)!}p^{m+1-0}(1-p)^0$$

$$= 1 - p^{n_k(g)}$$

And, as we have initially multiplied by (1-p)(m+1):

$$= \frac{1-p^{n_k(g)}}{(1-p)n_k(g)}$$

□

7.1.2 Proposition 1

To simplify the exposition $n_k(g)$ is abbreviated with n_k.

$$q(n_k) = \frac{1 - p^{n_k}}{(1-p)n_k}$$

$$\frac{\partial q(n_k)}{\partial n_k} = \frac{-p^{n_k} ln(p)(1-p)n_k - (1-p^{n_k})(1-p)}{(1-p)^2 n_k^2}$$

$$= \frac{-p^{n_k} ln(p) n_k - (1-p^{n_k})}{(1-p)n_k^2}$$

$$= \frac{p^{n_k} - p^{n_k} ln(p) n_k - 1}{(1-p)n_k^2}$$

$$= \frac{p^{n_k}(1 - ln(p)n_k) - 1}{(1-p)n_k^2} < 0 \ for \ p \in (0,1), \ n_k > 0$$

$\square$

7.1.3 Proposition 2

In what follows, $n_i(g)$ is abbreviated with n_i. Furthermore, we assume for simplicity that the network around i is regular, so that $q(n_j(g)) = \bar{q} \, \forall \, j \in N_i(g)$:

$$Q_i(g) = 1 - \prod_{j \in N_i(g)} (1 - pq(n_j(g))$$

$$Q_i(g) = 1 - (1 - p\bar{q})^{n_i}$$

$$\frac{\partial Q_i(g)}{\partial n_i} = -(1 - p\bar{q})^{n_i} ln(1 - p\bar{q})$$

Notice that $0 < 1 - p\bar{q} \leq 1$ as $\bar{q} \in [0,1]$ and $p \in (0,1)$. This implies that $ln(1 - p\bar{q}) \leq 0$ and that $(1 - p\bar{q})^{n_i} > 0$. As the whole expression is multiplied by -1, this shows that $\frac{\partial Q_i(g)}{\partial n_i} \geq 0$, which proves the first part of proposition 2.

$$\frac{\partial^2 Q_i(g)}{\partial n_i^2} = -(1 - p\bar{q})^{n_i} (ln(1 - p\bar{q}))^2$$

Now $(ln(1 - p\bar{q}))^2 \geq 0$ and $(1 - p\bar{q})^{n_i} > 0$. Hence $\frac{\partial^2 Q_i(g)}{\partial n_i^2} \leq 0$. $\square$

7.2 Figures

Figure 2: The job contact network of a block in urban Ethiopia

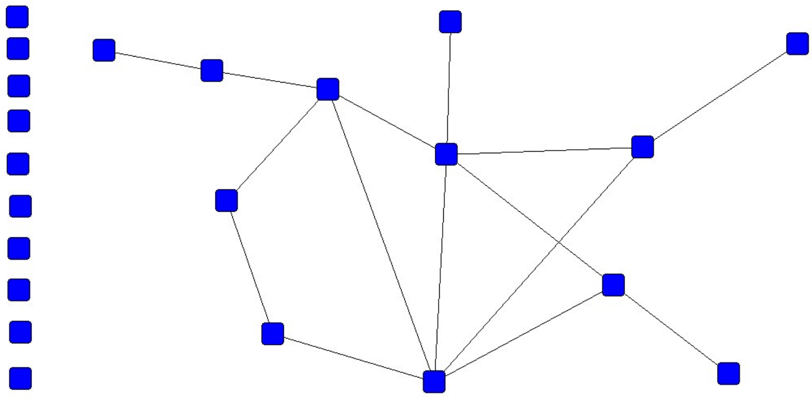


Figure 3: Distribution of links in job contact networks

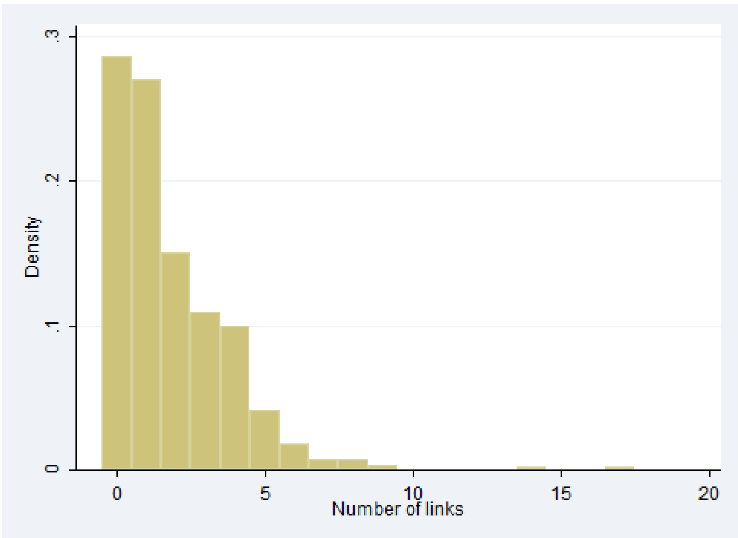


Figure 4: Number of days i has talked to j in the last 30 days, excluding dyads where i does not know j

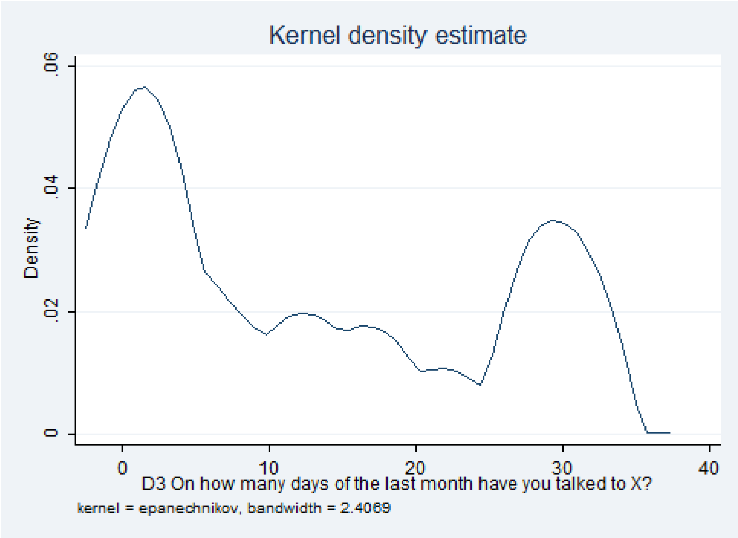


Figure 5: Relation between i and j

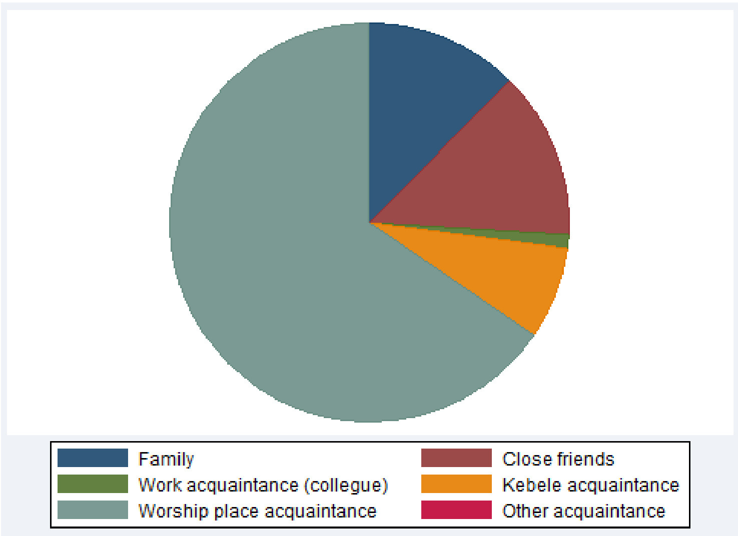


Figure 6: Results from Dictator Game

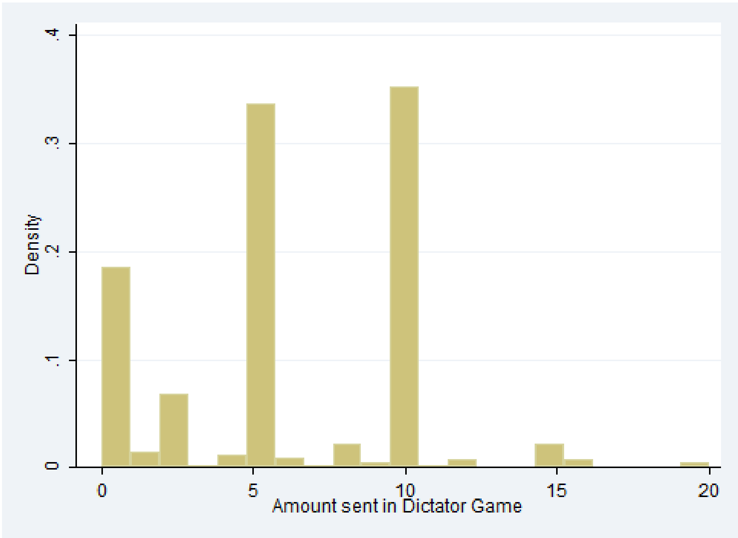


Figure 7: Self-reported motives behind linking decisions in the SELFa and SELFa2 treatments

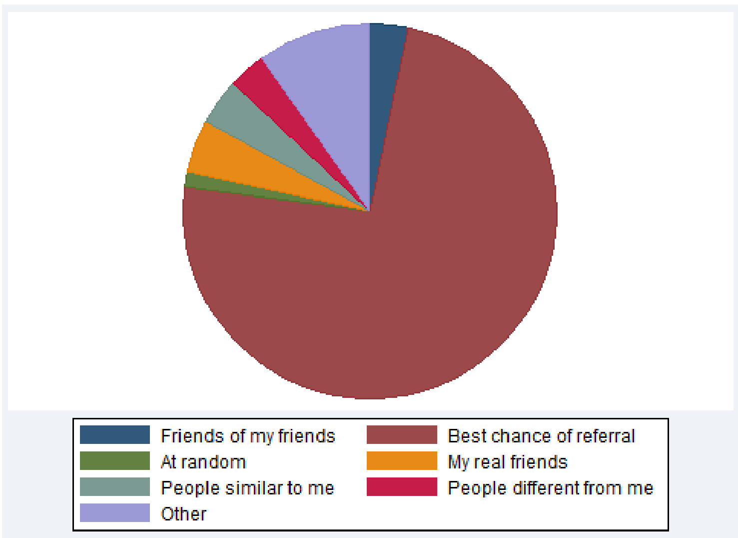
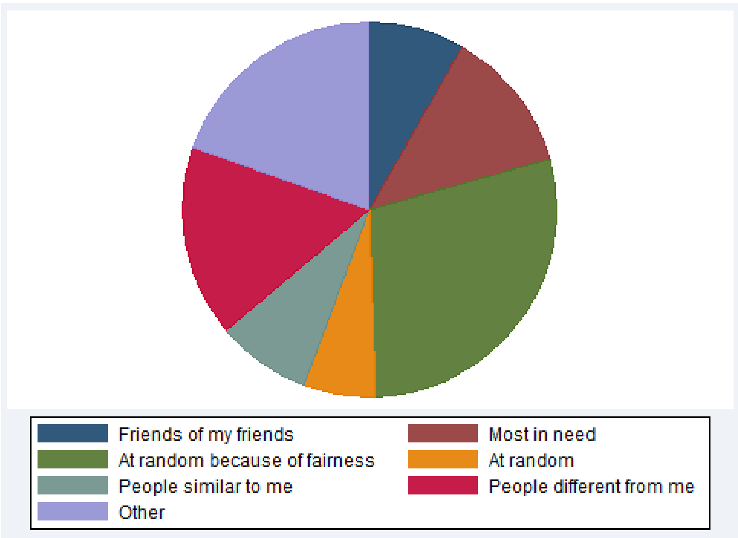


Figure 8: Self-reported motives behind linking decisions in the OTHERa treatment



7.3 Tables

Table 4: OLS Regression: Selection into the experiment

	Male	Age	Education	Muslim	Migrant	Employed	Inactive	Earnings	Degree	Degree2
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Selected	-.065	-1.394	.876	-.226	-.041	-.041	.014	-146.037	1.426	.527
	(.062)	(.885)	(.605)	(.050)***	(.058)	(.060)	(.051)	(154.368)	(.400)***	(.509)
Obs.	496	496	501	495	488	501	501	501	501	503

OLS regression. Column headings indicate the dependent variable. Degree captures to the self-reported number of friends in the block. Degree2 captures to the number of residents of the blocks who report i as their friend. Confidence: *** ↔ 99%, ** ↔ 95%, * ↔ 90%. Standard errors corrected for clustering at residential block level are reported in parenthesis.

Table 5: OLS Regression: Covariates balance across network centrality

	Male	Age	Education	Muslim	Migrant	Employed	Inactive	Earnings	Degree
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Centrality j=2	.056	1.313	-.524	-.072	.017	.047	-.009	103.861	-.405
	(.045)	(1.371)	(.600)	(.047)	(.051)	(.051)	(.067)	(93.997)	(.338)
Centrality j=3	-.006	.751	-.379	-.054	.022	.067	-.025	-32.703	.335
	(.050)	(.973)	(.548)	(.058)	(.051)	(.043)	(.056)	(85.064)	(.255)
Obs.	430	430	430	429	421	430	430	430	430

OLS regression. Column headings indicate the dependent variable. Degree captures to the self-reported number of friends in the block. Confidence: *** ↔ 99%, ** ↔ 95%, * ↔ 90%. Standard errors corrected for clustering at session level are reported in parenthesis.

Table 6: OLS Regression: Covariates balance across SELF and OTHER treatments

	Male	Age	Education	Muslim	Migrant	Employed	Inactive	Earnings	Degree
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
SELF Treatments	.087	1.198	.342	-.041	-.010	-.043	.069	47.185	-.431
	(.044)**	(.557)**	(.609)	(.069)	(.053)	(.071)	(.073)	(85.801)	(.236)*
Obs.	430	430	430	429	421	430	430	430	430

OLS regression. Column headings indicate the dependent variable. Degree captures to the self-reported number of friends in the block. Confidence: *** ↔ 99%, ** ↔ 95%, * ↔ 90%. Standard errors corrected for clustering at session level are reported in parenthesis.

Table 7: Linear Probability Model: Understanding

	Understanding1	Understanding2
	(1)	(2)
SELF non anonymous	.092	
	(.078)	
OTHER	.019	
	(.085)	
OTHER non anonymous	-.180	
	(.133)	
SELF control for priming	-.039	-.030
	(.101)	(.081)
Const.	.800	.792
	(.070)***	(.037)***
Obs.	427	427

OLS regression. Dependent variable is a dummy which takes a value of one if the respondent answered correctly both questions about the incentives of the game. Confidence: *** ↔ 99%, ** ↔ 95%, * ↔ 90%. Standard errors corrected for clustering at session level are reported in parenthesis.

Table 8: Summary of i's experimental decisions by treatment and j's centrality

	Treatments											
j is...	SELFa		SELFn		SELFa2		OTHERa		OTHERn		Total	
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
Centrality 1	84.0	49.4	98.0	47.6	59.0	43.4	71.0	37.8	54.0	33.8	366.0	42.6
Centrality 2	42.0	24.7	54.0	26.2	31.0	22.8	39.0	20.7	47.0	29.4	213.0	24.8
Centrality 3	19.0	11.2	39.0	18.9	29.0	21.3	50.0	26.6	45.0	28.1	182.0	21.2
Random	25.0	14.7	15.0	7.3	17.0	12.5	28.0	14.9	14.0	8.8	99.0	11.5

Table 9: Linear Probability Model: SELF treatments

	Base	Controls	Treatments
	(1)	(2)	(3)
j centrality = 2	-.167	-.179	-.195
	(.001)***	(.010)***	(.002)***
j centrality = 3	-.198	-.200	-.340
	(.001)***	(.016)**	(.012)**
Non anonymous			-.023
			(.650)
Non anonymous X c = 2			.020
			(.758)
Non anonymous X c = 3			.160
			(.050)**
Control for priming			-.035
			(.726)
No priming X c = 2			.005
			(.970)
No priming X c = 3			.038
			(.832)
Const.	.397	.407	.436
	(.000)***	(.000)***	(.000)***
Obs.	1594	1528	1528

Dyadic OLS regression. Dependent variable is a dummy which takes a value of one if i chose to establish a link with j. Columns 2 and 3 include controls for gender, religion and earnings and for the interaction between these variables and the dummies for the j's centrality. Confidence: *** $\leftrightarrow$ 99%, ** $\leftrightarrow$ 95%, * $\leftrightarrow$ 90%. Standard errors corrected for clustering at session level. P-values obtained with wild bootstrap-t procedure reported in parentheses.

Table 10: Linear Probability Model: SELF treatments

	Understanding1	Understanding2	OtherRegarding
	(1)	(2)	(3)
j centrality = 2		-.016	-.159
		(.904)	(.196)
j centrality = 3		-.010	-.249
		(.942)	(.044)**
Understanding	.131	.125	
	(.000)***	(.032)**	
Understand X c = 2	-.199	-.187	
	(.000)***	(.140)	
Understand X c = 3	-.229	-.222	
	(.004)***	(.038)**	
DG sent dummy			-.028
			(.706)
Sent dummy X c = 2			-.022
			(.824)
Sent dummy X c = 3			.055
			(.604)
Const.	.290	.298	.431
	(.000)***	(.000)***	(.000)***
Obs.	1517	1517	1528

Dyadic OLS regression. Dependent variable is a dummy which takes a value of one if i chose to establish a link with j. All columns include controls for gender, religion and earnings and for the interaction between these variables and the dummies for the j's centrality. Confidence: *** ↔ 99%, ** ↔ 95%, * ↔ 90%. Standard errors corrected for clustering at session level. P-values obtained with wild bootstrap-t procedure reported in parentheses.

Table 11: Linear Probability Model: OTHER treatments

	Base	Controls	Treatments
	(1)	(2)	(3)
j centrality = 2	-.065	-.099	-.153
	(.308)	(.162)	(.170)
j centrality = 3	.012	-.082	-.123
	(.884)	(.194)	(.340)
Non anonymous			-.029
			(.996)
Non anonymous X c = 2			.112
			(.328)
Non anonymous X c = 3			.082
			(.608)
Const.	.302	.334	.349
	(.000)***	(.000)***	(.000)***
Obs.	1072	1028	1028

Dyadic OLS regression. Dependent variable is a dummy which takes a value of one if i chose to establish a link with j. Columns 2 and 3 include controls for gender, religion and earnings and for the interaction between these variables and the dummies for the j's centrality. Confidence: *** ↔ 99%, ** ↔ 95%, * ↔ 90%. Standard errors corrected for clustering at session level. P-values obtained with wild bootstrap-t procedure reported in parentheses.

Table 12: Linear Probability Model: OTHER treatments

	Understanding	OtherRegarding	Interaction
	(1)	(2)	(3)
Understanding	.032		.007
	(.516)		(.990)
Understand X c = 2	-.042		
	(.540)		
Understand X c = 3	-.018		
	(.636)		
DG sent dummy		-.019	-.014
		(.826)	(.796)
Sent dummy X c = 2		-.015	
		(.856)	
Sent dummy X c = 3		.051	
		(.620)	
Understand x sent dummy			-.039
			(.668)
Und X Sent dummy X c = 2			.064
			(.410)
Und X Sent dummy X c = 3			.088
			(.316)
Const.	.269	.285	.279
	(.000)***	(.000)***	(.000)***
Obs.	1021	1028	1021

Dyadic OLS regression. Dependent variable is a dummy which takes a value of one if i chose to establish a link with j. All columns include controls for gender, religion and earnings and for the interaction between these variables and the dummies for the j's centrality. Confidence: *** ↔ 99%, ** ↔ 95%, * ↔ 90%. Standard errors corrected for clustering at session level. P-values obtained with wild bootstrap-t procedure reported in parentheses.

Table 13: Linear Probability Model: Non anonymous treatments

	SELFn	OTHERn
	(1)	(2)
j centrality = 2	-.177	.002
	(.002)***	(.584)
j centrality = 3	-.165	.058
	(.004)***	(.999)
i knows j	.101	.241
	(.094)*	(.126)
Same gender	-.055	-.003
	(.046)**	(.910)
Same migrant status	.039	.051
	(.520)	(.212)
Both Muslim	.024	.044
	(.540)	(.340)
Sum age	.002	-.002
	(.458)	(.146)
Diff age	-.002	.003
	(.092)*	(.088)*
Const.	.303	.320
	(.000)***	(.000)***
Obs.	563	452

Dyadic OLS regression. Dependent variable is a dummy which takes a value of one if i chose to establish a link with j. Confidence: *** $\leftrightarrow$ 99%, ** $\leftrightarrow$ 95%, * $\leftrightarrow$ 90%. Standard errors corrected for clustering at session level. P-values obtained with wild bootstrap-t procedure reported in parentheses.