I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 18258

Why Do Workers Make Job Referrals? Experimental Evidence from Ethiopia

Marc Witte

NOVEMBER 2025

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 18258

Why Do Workers Make Job Referrals? Experimental Evidence from Ethiopia

Marc Witte
VU Amsterdam, Tinbergen Institute and IZA

NOVEMBER 2025

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

IZA DP No. 18258 NOVEMBER 2025

ABSTRACT

Why Do Workers Make Job Referrals? Experimental Evidence from Ethiopia*

What motivates workers' referral decisions? Combining a field experiment in a firm and urban social network data, I first show that workers primarily refer those who previously referred them. This reciprocity leads to significant on-the-job productivity losses and excludes less connected individuals. Incentivized referrals reduce reciprocity and make workers screen more productive colleagues. Second, peripheral workers use referrals strategically to establish new and reciprocated links which persist after 18 months. These results are consistent with a network-based referral model where individuals trade off pecuniary and social incentives. The findings suggest that referrals through social networks can reinforce labor market inequalities.

Keywords: job referrals, social networks, field experiment

Corresponding author:
Marc Witte
VU Amsterdam
School of Business and Economics
De Boelelaan 1105
1081 HV Amsterdam
The Netherlands
E-mail: m.j.witte@vu.nl

* I am grateful to Girum Abebe, Johannes Abeler, Steffen Altmann, Francis Bloch, Stefano Caria, Marcel Fafchamps, James Fenske, Simon Franklin, Douglas Gollin, Morgan Hardy, David Huffman, Clément Imbert, Ingo Isphording, Pramila Krishnan, Sylvie Lambert, Julien Labonne, Simon Quinn, Christopher Roth, Margaret Stevens, Alemayehu Seyoum Taffesse, Christopher Woodruff and Liam Wren-Lewis for helpful comments and to Alemayehu Woldu and Mersha Beletew for outstanding research assistance. I received helpful feedback from seminar participants at Aix-Marseille, Bilkent, Birmingham, Collegio Carlo Alberto, Exeter, Gothenburg, IZA, NYUAD, Oxford, UC3M and UEA, and from conference participants at NEUDC, SEEDEC, ABCA, RES, Warwick, CSAE, EEA, TADC-LBS and GLAD. Data collection and experimental implementation were funded by the International Growth Centre. The project would not have been possible without the constant support of Rose Page, Simon Quinn, CSAE, and EDRI. The research described in this article was approved by the Economics Departmental Research Ethics Committees (DREC) at the University of Oxford (ECONCIA16-17-003). This field experiment was registered in the American Economic Association Registry for randomized control trials under trial number AEARCTR-0002334.

1 Introduction

Job referrals in social networks are ubiquitous: across countries a substantial share of workers obtains their job through social contacts. This is even more important in developing countries—with stronger spatial and informational frictions and more informality—where up to 72% of workers state that they found their job through networks (Sapin et al. (2020); appendix figure A3).[1] In principle, referrals can be mutually beneficial for employers (e.g. through improved peer monitoring or reduced screening costs) and job seekers (e.g. through reduced search costs) alike.[2] Little is known, however, about one of the most important parties in the referral process: the incumbent worker who *makes* the referral. We do not know which considerations influence the referral decisions of workers, and how these interact with the institutional features of the referral process—even though they directly affect who gets invited to the job and ultimately how the firm performs. This is an important gap: workers can generally decide between multiple potential referrals, which means that we can expect workers' referral strategies to affect the workforce composition of networked firms. Similarly, worker referrals might be a source of persistent inequality in the access to labor markets, as individuals compete for limited referrals and some people will be left out. Finally, the prevalence of referrals renders them suitable as a potential policy lever.

This paper investigates the motivations behind job referrals by experimentally varying the monetary and non-monetary benefits for workers referring other people

[1] Studies investigating social networks and labor market outcomes include Hellerstein et al. (2011), Topa (2011), Kramarz and Skans (2014), Schmutte (2015) for developed and Wahba and Zenou (2005), Munshi and Rosenzweig (2006), Beaman and Magruder (2012), Heath (2018) for developing countries.

[2] Empirically the effects of referrals on job-seekers and employers have been investigated, with mixed results. First, referrals reduce search costs for job-seekers (Calvó-Armengol, 2004, Galeotti and Merlino, 2014). Second, from the perspective of the employer, relying on worker referrals can improve peer monitoring (Kugler, 2003, Heath, 2018) and the screening of productive workers (Montgomery, 1991, Munshi, 2003, Beaman and Magruder, 2012). Third, there is growing empirical evidence that not everybody is well served by referrals (e.g. Beaman et al., 2018, Chandrasekhar et al., 2020).

to a job. To do this, I set up a data entry firm with over 500 employees in Addis Ababa, Ethiopia, where workers can refer other people from their social network for the job. I randomly vary the conditions of this job referral process along two dimensions. The first dimension is a monetary incentive payment: some workers receive a performance incentive according to the work performance of the individual they refer, while others do not get incentivized. The second dimension is the visibility of the referral, i.e. the degree to which the referred contact is aware of who invited them. Referrals are either completely open (taking place in the field), partly anonymous (where invited workers can reciprocate the referral, but do not know the referee) or completely anonymous (no reciprocity possible). These two treatment variations enable me to cleanly identify how workers trade off monetary and non-monetary payoffs and altruistic considerations while referring others for jobs. The experiments generates a panel of workers' referral decisions over multiple work days, which allows me to account for mechanisms that are specific to repeated (referral) interactions.

Three central empirical challenges emerge when studying the motivations behind job referrals. First, understanding the choice set that workers face when making a referral decision requires understanding the pool of potential referrals, which can be infinite in theory. I circumvent this problem by collecting complete and detailed social network data from 16 urban neighborhoods in which the workers live, for a total of 739 individuals. This includes comprehensive dyadic information on the workers' social connections within the neighborhood as well as aggregate information on out-of-neighborhood contacts, allowing me to provide clean evidence on the type and intensity of the connection between the worker and her referral. Second, networks evolve endogenously. Several recent studies investigating the impact of social connections on labor market outcomes overcome this issue by providing exogenous shocks to a network, e.g. by introducing a set of individuals to one another and inducing them to interact for some period.[3] This paper takes a different approach, by letting the individual rather than the researcher choose whom to connect with, but exogenously varying the conditions of referral. Finally, various

[3]For example Feigenberg et al. (2013), Fafchamps and Quinn (2016), Cai and Szeidl (2017), Abebe et al. (2020), Vasilaky and Leonard (2018).

considerations simultaneously drive referral decisions observed in reality, complicating the identification of strategic motives in the process. By randomly varying components of the referral payoff, my experiment overcomes this challenge and allows me to cleanly disentangle the mechanisms at play.

In the field experiment, I exogenously vary two key dimensions of the referral process: referral incentives and referral visibility. I study a total of three treatments and a control condition (see table 1), which are informed by a simple theoretical model relating a job referral's direct performance payoff (through monetary incentives), its social payoff (including long-run considerations) and the productivity of the referred worker. The control case is an un-incentivized open job referral, mirroring how most referrals work in reality. I compare the other treatments against this condition. In the first margin of variation, I introduce an incentive treatment, meaning that the worker making a referral gets a linear financial reward based on the performance of the invited worker. The second margin of variation (applied only to un-incentivized referrals) is the visibility of the referrals, in which I introduce partly anonymous or fully anonymous referrals. To ensure anonymity, these referrals are made to individuals outside the worker's own neighborhood of residence. Partly anonymous referrals only allow the invited worker to reciprocate the job invitation without knowing who invited them. In fully anonymous referrals, the invited worker does not know who invited them and cannot reciprocate. The experiment is conducted over three rounds (or work stints), with two rounds of job referral decisions after work. This panel of worker productivity and referral data enables me to analyse repeated interactions.

This paper contains four key findings, which can be interpreted through the lens of my simple theoretical framework. First, the job invitations are characterized by a substantial degree of reciprocal referring when conducted openly and without referee incentives. A casual worker who invites another neighborhood contact to the day job is almost 30 percentage points more likely to be re-invited by this same person, compared to being invited by a random worker. Importantly, these return invitations come at a substantial cost for productivity. Workers referred with a return invitation produce between five and ten percent less output than non-reciprocally invited individuals, despite always being paid the same piece rate for each produced unit. This suggests that workers receive higher social payoffs from handing out return invitations (e.g. through social rewards or future benefits), which prevents them from referring the most productive workers to the job. The high level of referral reciprocity also suggests that individuals with one-off referral opportunities have higher chances of receiving work in the future.

Second, reciprocity is substantially reduced under the incentivized referral treatment, where the number of return invitations decreases by approximately 20%. This demonstrates the trade-off between direct performance payoffs and social (including long-run strategic) payoffs, as shown in my model. I also find that when given the performance incentive, workers refer more productive individuals to the job, a finding that is consistent with Beaman and Magruder (2012). If the other worker is invited with an incentivized referral, she performs approximately 18% better on the job, compared to workers referred without incentives. I provide additional evidence that this result reflects the referred worker's innate productivity and is not due to some behavioral adjustment such increased learning or motivation—ultimately demonstrating that workers are able to screen more productive individuals in their local, relatively homogenous neighborhood labor markets.

My third finding exploits the exogenous variation in the visibility of the referral. I find that under the standard open referral treatment, workers refer socially well-connected individuals, as measured by their social network centrality.[4] This is in line with my model, where referrals are not just made with the immediate expectation of direct reciprocity, but also with regard to future network benefits, such as access to information.[5] As a consequence, individuals without many connections are left jobless. This referral behavior cannot be explained by socially connected individuals being more productive workers: I find no significant correlation between an individual's network centrality and her on-the-job productivity. Under the referral treatments which are partly or fully anonymous, less central but equally

[4] I show that this is over and above the level that can be expected from the 'friendship paradox', i.e. the phenomenon that one's friends tend to be more central in a network (see Eom and Jo, 2014).

[5] This is the opposite mechanisms as in Bala and Goyal (2000), who predict star-shaped networks (with all connections leading to the central node) in the case of bilaterally flowing non-rival goods such as information.

productive workers are selected. Two different mechanisms explain this: Under partly anonymous referrals, immediate (and rival) referral reciprocity is the only benefit that matters to the employee, which is maximised by inviting non-central referrals. Under fully anonymous referrals, any network benefits are eliminated and only other-regarding preferences matter: here, the least central workers are chosen altruistically, out of sympathy with the socially excluded.

Lastly, I show that labor markets relying on job referrals can persistently exclude individuals from jobs. Approximately 9% of respondents report not knowing any other respondent in their area, and a quarter of respondents only know up to two people.[6] These peripheral individuals are rarely referred by other workers and thus remain excluded. In contrast, peripheral individuals who are (randomly) selected to work and refer others use these opportunities as a device to connect to larger components of the social network. They are—non-mechanically—more than twice as likely to invite a previously unknown worker to the job than individuals who are already well-connected at baseline. These new work links are neither unilateral nor short-lived. On the contrary, the other worker, now a new link in the peripheral worker's network, is more likely to reciprocate and provide the peripheral worker with a return job, compared to a connection that already existed at baseline. 18 months after the experiment, almost 40% of individuals who did not know each other at baseline are still in contact, demonstrating that a very light-touch intervention can have long-run effects on connectedness in local labor markets.

My paper contributes to four different strands of literature. First of all, I speak to the literature on the costs and benefits of job referrals. A number of papers have demonstrated how referrals can reduce search costs for job-seekers (Calvó-Armengol, 2004, Galeotti and Merlino, 2014) or solve moral hazard (Kugler, 2003, Heath, 2018, Dhillon et al., 2020) and information asymmetry issues (Montgomery, 1991, Munshi, 2003, Beaman and Magruder, 2012, Dustmann et al., 2016) for employers. Here, I focus on the motivations of the agent, i.e. the worker making the referral, and how she benefits from the act of referring someone to become her coworker. In this literature, my paper relates most closely to Beaman and Magruder

[6]For these peripheral nodes, the average number of all social contacts *including* individuals living outside the neighborhood is four, and hence equally low.

(2012). In this field experiment, the authors exogenously vary the monetary incentives of a one-off job referral process and find that under incentives participants refer high-ability workers to tasks (in line with what I find). I complement Beaman and Magruder (2012)'s study in at least two important ways: first, my study design allows me to observe repeated rounds of referral decisions, showing that reciprocity is an important motivation for referrals. Since the dynamic nature of reciprocal referrals increases workers' social pay-offs and leads to lower work performance of the invited worker, I can show that the longer-run detrimental consequences of reciprocal referrals warrant incentive pay, even more than a one-shot perspective would suggest. Second, I add the dimension of labor market exclusion to Beaman and Magruder (2012)'s work. By collecting complete network data prior to the experiment, I can relate the job referral decisions to the full, real-life social network of the workers and describe how peripheral workers strategically enlarge their network.

Second, I contribute to the broader literature on the role of networks in developing countries, in particular in labor markets.[7] Referrals to jobs illustrate the broader phenomenon that social networks shape labor markets and individual outcomes. People help each other find employment (Granovetter, 1973, Ioannides and Datcher Loury, 2004), and individuals are more likely to be employed if their friends are (Topa, 2001, Calvó-Armengol and Jackson, 2004). In contrast to a recent literature exogenously shocking the labor market network of interest (Feigenberg et al., 2013, Fafchamps and Quinn, 2016, Cai and Szeidl, 2017, Vasilaky and Leonard, 2018, Abebe et al., 2020), I choose an approach in which the network members act themselves, under exogenously varying constraints (similar to Beaman and Magruder (2012) and Beaman et al. (2018)).

Third, despite record levels of urbanisation, social networks in urban areas, particularly in developing countries, are relatively understudied. Most empirical work either focusses on village communities (e.g. Cai et al., 2015, Chandrasekhar et al., 2018, Banerjee et al., 2019, Breza and Chandrasekhar, 2019) or other social connections, such as family networks (e.g. Magruder, 2010), where geographic proximity or kinship mediate network effects. This is an important gap in the literature, as we know from mostly descriptive studies that employment networks are often

[7] Chuang and Schechter (2015) provide an extensive summary of the literature.

situated in urban neighborhoods (Hipp et al., 2012, Hellerstein et al., 2014, Caria et al., 2020). Similar to Rigol et al. (2020), I show that private information about social connections can be elicited in the context of urban communities. The workers in my sample display high levels of sophistication in the way they strategically use the referral opportunities, which contrasts markedly with the low sophistication of small firm owners in the same context (Abebe et al., 2020).

Lastly, this study relates to a smaller literature on strategic network formation, mostly in the context of laboratory studies. This literature shows that individuals decide to link with more central participants in situations where the social network structure is artificially imposed, either in laboratory games without direct human interactions (Conte et al., 2015) or by ascribing participants a random network position (Caria and Hassen, 2013). I provide the first empirical evidence that a similar strategic network formation process takes place within existing, real-life social networks, where links are formed 'in the field'. Moreover, in line with the foundational work in Fehr and Gächter (2000) and Gächter and Falk (2002), I show that considerations of reciprocity play an important role in job referral decisions.

This paper can inform a policy debate around labor market inequalities induced by social networks. I find that job referrals act as significant barriers to entry for socially detached individuals, which can lead to persistence in unemployment. However, as opposed to a recent literature emphasizing the potential costs of relying on social networks for hiring (Calvó-Armengol and Jackson, 2004, Beaman et al., 2018, Galenianos, 2020), I identify one-off job-referrals as a way of escaping the 'network poverty trap'. This finding has implications for how policymakers could think of alleviating long-term unemployment, along two margins. First (and passively), changing the incentive structure of existing workers making referrals can help the inclusion of isolated job-seekers. Second (and actively), long-term unemployed could be integrated by providing them with subsidised temporary jobs as linking opportunities, in particular in settings of high informality.

The remainder of this paper is structured as follows: in section 2, I develop a conceptual framework of job referrals in social networks. Section 3 presents the experimental design, while section 4 introduces the context, the implementation, and summary statistics. Section 5 lines out the empirical specifications. The empirical results are presented in section 6 and 7. Section 8 concludes.

2 Conceptual framework

In this section, I present a stylised model of job referrals in social networks. The model builds on the theoretical frameworks presented in Bandiera et al. (2009) and Beaman and Magruder (2012). I then extend the latter version by explicitly modelling the different payoffs individuals can receive in their social network.

2.1 Job referrals in social networks

Consider worker i, who has the opportunity to make one job referral to another person j from her neighborhood contacts. The decision of the worker is: whom should I refer to the job? The referral contract is specified by the employer. Each worker and referral can perform on the job with ability $\theta_i \in \{\theta_i^H, \theta_i^L\}$. For simplicity, ability is modelled as a binary that can take the values *high* (H) or *low* (L), so that there are two types of contacts, those of high ability and those of low ability.[8] The employer offers the worker a referral contract, consisting of an incentive payment (P_i) based on the referred contact j's performance. In particular, P_i is the amount that i gets if j produces output, and P_i is linear in j's output, i.e. $P_i = P_i(\theta_j) = p_i\theta_j$. I assume worker i can directly observe her contact's ability $\theta_j \in \{\theta_j^H, \theta_j^L\}$.[9]

Worker i might expect monetary pay-offs from the employer for referring contact j. The payoffs are a function of the contract type P_i and the referral's ability θ_j. Workers also might expect a social payoff σ_{ij} from the contact j. The social payoff can include cash transfers from j to i—but also immediate and future network benefits, such as reciprocal job referrals or information access. Given the two stylised types of friends, the social payoffs are $\sigma_{i1} \in \arg\max_j(\sigma_{ij}|\theta_j = \theta_j^H)$ for the high-ability and $\sigma_{i2} \in \arg\max_j(\sigma_{ij}|\theta_j = \theta_j^L)$ for the low-ability type. When selecting the high-ability contact under an incentivized referral scheme, i will receive

[8]The framework with continuous ability is derived in appendix section B.

[9]Alternatively, every worker could only observe a signal of her contacts' ability, $\hat{\theta}_i \in \{\theta_i^H, \theta_i^L\}$ with the probability of accuracy β_i, where $P(\theta = \theta^H|\hat{\theta} = \theta^H, i) = P(\theta = \theta^L|\hat{\theta} = \theta^L, i) = \beta_i$, so that $\beta_i \in [0.5, 1]$. This case is presented in appendix section B. The implications of the model stay the same.

the payoff $P_i + \sigma_{i1}$, whereas for selecting the low-ability type she will receive σ_{i2}. Assuming that every worker i among her contacts c_i has a choice between the high-ability and the low-ability type, the worker compares the best high-ability with the best low-ability type. As a consequence, worker i will refer a high-ability type if

$$P_i > \sigma_{i2} - \sigma_{i1}. \tag{1}$$

In other words, worker i will refer the high-ability type if the performance pay is larger than the differences in social benefits. If the worker only has low-ability contacts, she will choose the contact who provides her with the highest social benefits σ_{ij}. So far, this model follows closely Beaman and Magruder (2012), which build on Bandiera et al. (2009). For ease of exposition, the framework proceeds with a continuous ability parameter $\theta_i \in [0, 1]$, which is derived in appendix section B.

2.2 Decomposing the social pay-off

Consider a worker i who has c_i contacts. The social payoff worker i receives from referring contact j can be modelled explicitly: $\sigma_{ij} = \pi_{ji} + b_{ji} + \omega_{ij} m_j$, where π_{ji} are direct monetary payments from j to i, b_{ji} are network benefits flowing from j to i, (such as reciprocal referrals, network goods, or information access) and m_j is altruism, expressed as j's income in i's utility with j-specific weight ω_{ij}. The network benefits b_{ji} in turn can be expressed as $b_{ji} = \rho_{ji} + \kappa_{ji}$, where ρ_{ji} are return referrals (direct reciprocity) from j to i, and κ_{ji} are further (including future) benefits of being connected to j, such as access to j's information. As a result, the full expression for the social payoff becomes $\sigma_{ij} = \pi_{ji} + \rho_{ji} + \kappa_{ji} + \omega_{ij} m_j$. I assume throughout that the direct monetary payments π_{ji} are zero, as I cannot observe them during my experiment. The full expression for the referral payoff with continuous ability is:

$$\Pi_{ij} = P_i(\theta_j) + \sigma_{ij} = P_i(\theta_j) + \rho_{ji} + \kappa_{ji} + \omega_{ij} m_j = p_i \theta_j + \rho_{ji} + \kappa_{ji} + \omega_{ij} m_j \tag{2}$$

where the last step assumes that P_i is linear in j's output. Equation (2) shows that the social payoff σ_{ij} and the performance incentive P_i are substitutes. In particular, referral contracts that include an incentive should *ceteris paribus* make workers i select referrals j with lower social payoffs σ_{ij}, through lower network benefits b_{ji}, a lower utility weight ω_{ij}, or both.

2.3 The role of network centrality

Every worker i has c_i connections in her local neighborhood network. She also observes the number of connections of her contacts, c_j.[10] For simplicity, I interpret a worker i's number of connections c_i as her degree centrality.[11] A contact j's network centrality is going to affect the social payoff σ_{ij} through different channels: $\sigma_{ij} = \rho_{ji}(c_j) + \kappa_{ji}(c_j) + \omega_{ij} m_j(c_j)$.

First, the likelihood of worker i receiving a direct reciprocal referral from contact j *ceteris paribus* decreases with her centrality c_j, so that $\frac{\partial \rho_{ji}}{\partial c_j} < 0$. This is because a contact with fewer connections c_j is more likely to select i for a return referral. Second, the likelihood of receiving future benefits from being connected to contact j increase with j's centrality, so that $\frac{\partial \kappa_{ji}}{\partial c_j} > 0$. Intuitively, this is because a central contact has access to more network goods, which increases the likelihood that they get shared with i.[12] As a consequence, the relationship between all network benefits and a referral's network centrality is unclear, and depends on whether the direct reciprocity or future benefits dominate:

$$\frac{\partial b_{ji}}{\partial c_j} > 0 \text{ if } \frac{\partial \kappa_{ji}}{\partial c_j} > -\frac{\partial \rho_{ji}}{\partial c_j}. \tag{3}$$

When network benefits $b_{ji} = 0$, the social payoff reduces to other-regarding preferences: $\sigma_{ij} = \omega_{ij} m_j$. In line with models of social preferences and inequality aversion (Fehr and Schmidt, 1999, Bolton and Ockenfels, 2000, Charness and Rabin, 2002), I assume that the weight ω_{ij} with which j's income m_j is included in i's utility is decreasing in j's centrality: $\frac{\partial \omega_{ij}}{\partial c_j} < 0$. Thus, under only other-regarding preferences, the social pay-off is maximised by referring the least central workers to the job:

$$\frac{\partial \sigma_{ij}}{\partial c_j} < 0. \tag{4}$$

As a consequence, I can write the full referral payoff function as:

$$\Pi_{ij} = p_i \theta_j + \sigma_{ij} = p_i \theta_j + \rho_{ji}(\overset{-}{c}_j) + \kappa_{ji}(\overset{+}{c}_j) + \omega_{ij}(\overset{-}{c}_j) m_j \tag{5}$$

[10]One could assume that i observes a signal $\hat{c}_j$ of c_j, but I omit this for simplicity.
[11]The implications of the model are the same for most other centrality measures.
[12]This is certainly the case for non-rival goods such as information, while the implications for rival goods are more complicated, see Bala and Goyal (2000).

Finally, I assume for simplicity that the elements of the social payoff function only depend on j through c_j. In that case, I can rewrite equation (5) as:

$$\Pi_{ij} = p_i\theta_j + \rho_i(\bar{c}_j) + \kappa_i(\overset{+}{c}_j) + \omega_i(\bar{c}_j)m_j \tag{6}$$

As experimenter, I vary several parameters of this model. First, I randomly introduce the incentive payment P_i that worker i can receive for her referral's performance.[13] Second, I exogenously vary the network benefits b_{ji} and their components, by enabling partly anonymous and anonymous referrals to workers from other neighborhoods.[14] I can deduct several propositions from the conceptual framework that inform the empirical part of this paper:

Prop. 1: A rise in P_i increases the payoff of referring a high-ability type (eq. (1)).

Prop. 2: Performance payoffs P_i and social payoffs σ_{ji} are substitutes in determining the likelihood of a worker being referred (eq. (2)).

An increase in P_i will lower the social payoffs required from referral j for the same overall referral payoff Π_{ij}. This could be expressed through lower network benefits b_{ji}, a lower utility weight ω_i — or a combination of these.

Prop. 3: An increase of ρ_{ji} compared to κ_{ji} in the network benefits b_{ji} will lead to the selection of referrals with lower centrality (eq. (3)).

Prop. 4: When network benefits b_{ji} are switched off, workers will select referrals with the lowest centrality (eq. (4)).

3 Experimental design

I conduct a field experiment to investigate what motivates workers' job referral decisions. In this section, I present the design of the experiment, which is informed by the conceptual framework presented previously and directly tests its predictions.

In short, the design of the experiment follows three principal steps. In a first step, I create a casual day labor market: I randomly invite young workers and job

[13]This is similar to Beaman and Magruder (2012).

[14]While the partly anonymous referrals to other neighborhoods decrease the likelihood of future benefits, κ_{ji}, and encourage direct return referrals, ρ_{ji}, the anonymous referrals set the complete expression b_{ji} to zero.

seekers from urban neighborhoods to work on a real and paid data entry job, over several consecutive work days (or rounds), which is not announced to the participants ex ante. Second, after every work day, workers are given the opportunity to refer individuals from their local social network to the same job the next day. Finally, the conditions of this 'referral treatment' vary randomly along two principal dimensions: 1) incentivized vs un-incentivized referrals and 2) in-person vs gradually anonymized referrals. In addition to that, some individuals are invited randomly as an additional comparison category. The data entry job performed by the workers relates to real and large payoffs (paid as piece rates), is conducted individually, and requires some endurance and cognitive effort.

The participants of the experiments are young urban dwellers aged 18 to 29 who are not in permanent employment or education in order to guarantee that the participants spend sufficient time in their local neighborhoods. More details on the sampling frame and participant characteristics can be found in section 4.

3.1 Treatment variations

The experimental structure is presented in figure A1. The work experiment is conducted over three consecutive work days. Each work day ends with a job referral decision, yielding an individual-level panel of productivity and referral data.

At first, I randomly invite two-thirds of all eligible individuals from the neighborhood to the first work session, while the remaining third does not get an invitation (for the first work session). Among those individuals working in the first session, half are randomly allocated to the control referral treatment, and the other half of participant workers get a treated referral opportunity. The treatment variations follow the treatment design displayed in table 1. The control referral treatment (C) is an un-incentivized, open job referral, similar to 'standard' job referrals mostly observed in reality. Workers in this group can invite someone from their neighborhood network to the next work session without incentives and 'in the field'.[15] Treatment variation 1 (T1) introduces an incentive treatment, meaning that the worker making a job referral gets a financial reward which is linear in the performance of

[15]Even though it is not enforced, everyone in the referral treatment group always invited another worker to the next session.

Table 1: Treatment matrix

		Visibility		
		Open	Partly anon.	Anonymous
Incentivisation	No	C	T2	T3
	Yes	T1	-	-

the invited worker. This treatment positively 'shocks' the monetary referral payoffs described in section 2. The second margin of variation shocks the visibility of the referral. Workers in the control referral condition can additionally make partly anonymous or fully anonymous referrals to other workers. Under partly anonymous referrals (T2), the invited worker does not know the identity of the inviting worker, but can reciprocate the referral, while under fully anonymous referrals (T3), reciprocation is now allowed. More details on the treatment variations follow below.

Every participant can only make one referral to someone in her neighborhood, excluding referring oneself. The second work session is then comprised by individuals referred by the first round workers. I then again randomly provide half of the second-round workers with a control referral and the other half with a treated referral. Accordingly, the third work session is again composed of referred workers.[16]

Open 'control' referrals (C)

This is the control condition against which the other treatments are compared. As figure A1 shows, a random half of the participant workers can make these open referrals after their work stints. Specifically, participants can refer any other eligible person from their own neighborhood of residence, i.e. anyone in their local social network.[17] Participants can only invite one other person.[18] The actual refer-

[16]In a subset of neighborhoods, I add a sub-treatment called 'no referral' treatment. Workers in this group cannot invite a person. Instead, in subsequent work sessions, a proportionate number of workers from the neighborhood are *randomly* invited to the work (labelled as 'invited by computer'). This sub-variation is introduced as an additional comparison category that avoids worker selection. I control for this minor design variation with neighborhood and session fixed effects.

[17]This includes same day co-workers and people who did not attend.

[18]As a consequence, the 'most popular' person in a neighborhood cannot receive referrals from all participants—only the first invitation counts. This potentially at-

ral process takes place after the work sessions, i.e. through personal interactions of workers "in the field". To fix information about who lives in the neighborhood, workers are given a list with their neighborhood's census, including the other residents' names, contact details and connections in the neighborhood.[19] This (signed) list serves as a ticket for the referral during the next session and ensures that real-life contact between the inviting worker and the referred worker takes place.[20]

Treatment variation: incentivisation of referral (T1)

This first treatment variation introduces a performance incentive: a random half of participants i receives an incentivized referral opportunity ($P_i > 0$), meaning that whatever their referral j earns during the next session is matched for them ($P_i = P_j$). Every other aspect of this referral treatment is identical to the open control referral (C), where the participant workers do not receive any payment related to the performance of the worker they invite (P_i=0). This treatment variation allows me to test proposition 1, i.e. whether the introduction of P_i increases the number of high-ability workers, as expressed by better job performance.

Treatment variation: visibility of referral (T2 & T3)

The two treatment variations in referral visibility shock the components of the social payoff σ_{ij} in the conceptual framework. The treatments are based on within-subject variation for workers in the unincentivised referral group (C) (see table 1). These two treatments vary the extent to which the referred worker can reciprocate and see who invited him/her. To prevent treatment contamination, i.e. that the referrals become public knowledge, referrals in treatment groups T2 and T3 are made to individuals living in other neighborhoods. These workers are not within the participant's own social neighborhood network, but live in identical economic circumstances, so that information on who made the anonymous referrals will not diffuse through the network.[21] The variation in referral visibility takes two separate

tenuates the parameter estimates for network centrality.

[19]In appendix section E.3, I show experimentally that my results are robust to whether the other residents' number of connections is mentioned on the referral lists or not, based on a 'shrouding' exercise.

[20]Referral lists can be found in the supplementary online materials.

[21]The within-subject design is chosen to maximise statistical power.

forms, the order of which is randomized over sessions:

1. **Partly anonymous referral in other neighborhoods (T2)**: Workers choose who to invite to the job from a list of individuals living in a different neighborhood (similar to the one the worker lives in). The list shows the potential referrals' age, sex and number of neighborhood connections (i.e. network centrality).[22] The salience of network centrality is thus constant across treatments. When the worker decides on who to refer from this list, her mobile phone number is passed on to the person she invites, and the invited worker is explicitly primed to reciprocate the inviting worker.[23] As with the other treatments, only one such reciprocal referral can be made. From a theoretical perspective, this treatment encourages immediate reciprocity ρ_{ji} concerns, whereas future network benefits κ_{ji} between the inviting and the invited worker cannot be enforced by the inviting worker.

2. **Anonymous referral in other neighborhoods (T3)**: This referral is similar to the partly anonymous referral in other neighborhoods, but no information about the inviting worker is passed on. The referral is therefore completely anonymous and reciprocity is not possible.

4 Context, implementation, and data

This study is situated in Addis Ababa, the capital of Ethiopia and large urban metropolis of over 5 million residents. Despite Ethiopia's strong economic growth in recent years, most of the country's young urban population is out of permanent or formal employment. Descriptive research has emphasised the importance of referral networks in the Ethiopian urban labor market (Serneels, 2007, Caria and Hassen, 2013), with a large share of both formal and informal sector jobs filled through job seekers' social networks. Socially disconnected individuals risk isolation from job information and referral sources, which is exacerbated by the fact

[22]The individuals listed are real and similar to the original workers in terms of sociodemographics. The referral choice of the original experiment workers is implemented in reality.

[23]These interactions are facilitated by the research team. The reciprocation prompt means that the invited worker is explicitly encouraged to 'thank' the inviter by selecting her for a similar job. Both the inviting and the invited worker know about these details.

that labor market information in Addis Ababa is exceptionally centralised and geographically restricted—most jobs are only publicly advertised on centrally located vacancy boards (Franklin, 2016, Abebe et al., 2021).[24]

This paper focusses on local neighborhood networks of young casual workers and job-seekers. Local information exchange and referrals to casual jobs are dominant in informal labor markets in developing countries (Wahba and Zenou, 2005, Serneels, 2007, for Ethiopia). To illustrate this: in my baseline sample, 95% of those individuals who ever approached another person for help in getting a job asked in their neighborhood, almost always involving an exchange of relevant information. In 40% of cases, the exchange of job information led to a direct referral to a job. Individuals out of permanent education or employment often obtain actual work and information about work through these neighborhood networks.

4.1 Neighborhood and participant sampling frames

In this subsection, I describe the sampling of neighborhoods within the city and the sampling of participants within neighborhoods. In a first step, I randomly select 16 urban neighborhoods in Addis Ababa, from a sample of over 2000 eligible ones.[25] The neighborhoods included in the randomisation sample represent the 'average' urban neighborhood in Sub-Saharan Africa—densely populated, with mostly single-storied buildings without any compounds or gardens. More than 65% of the population of Addis Ababa live in such neighborhoods.[26]

Within the selected neighborhoods, a complete census (via door-to-door solicitation[27]) of all eligible resident individuals is conducted, where eligible individuals 1)

[24]Internet penetration in Ethiopia overall was at 11.5% in 2015, but substantially higher in Addis Ababa (The World Bank, 2016). However, the few online or mobile phone services for job-seekers that have been established recently are still very small in size, both in terms of users as well as job vacancies on offer.

[25]A neighborhood is defined as an enumeration area from the Addis Ababa census, in line with the literature on urban networks (Bayer et al., 2008).

[26]The remaining 35% are equally divided between multi-story condominium complexes, compound / gated housing, and informal shantytowns in the outskirts.

[27]In practice, this means that enumerators went to every household in the neighborhood and asked if an individual satisfying the eligibility criteria lives in the household. If yes, the eligible individual's name as well as basic demographic and

permanently live in the selected neighborhood, 2) are between 18 and 29 years old, and 3) are not in permanent employment or education.[28] About 80% of the population within the chosen age group satisfy these criteria, indicating that the sample represents a substantial fraction of the urban labor market. These eligible individuals are then surveyed, including a detailed social network section. The final sample consists of 739 individuals from 16 urban neighborhoods who were surveyed during baseline in November/December 2016; for summary statistics see section 4.3. Additional details on sampling can be found in the supplementary online materials.

4.2 Data entry job description and work logistics

After completion of the baseline survey, participants are invited to the experimental work session, separated by neighborhood, which is announced to them as a paid 'day job'. Upon arrival at the work site (usually a rented public hall close to the respective neighborhood), a random two-thirds of present individuals are selected to participate in the first work session.[29] The ensuing data entry job requires some effort from the workers: participants are handed employee data from a company that they have to sort by employee work team and age. Similar tasks has been tested in the same context (Abebe et al., 2021).[30] Each individual's work is compensated with a piece rate of ten ETB per correctly entered work team, for a maximum of 20 teams / 200 ETB ($\simeq$ 12.00 USD). This is a non-negligible payment compared to the sample's median weekly spending of 300 ETB. After completing the task, participants are asked to make the different job referrals explained previously. The work sessions take place on consecutive days or are spread over a whole day, depending on logistical factors. In the empirical analysis, I include session fixed effects to control for any heterogeneity caused by different sessions. The second

contact details were recorded, to populate the baseline network roster.

[28]Permanent employment is defined as full-time employment at a single employer, including self-employment, going at least one month back. Permanent education is a full-time degree course at a university, school or training institute.

[29]The other people are reimbursed for transportation expenses and leave.

[30]It was not possible to conduct physical tasks, which would be closer to the reality of casual workers in this context (e.g. on construction sites). However, the work only depends on individual performance, demands real effort, and does not contain any elements of service-style or customer-facing jobs.

and third sessions of the experiment are unannounced to the participants ex ante. It is possible that participants expect a third session once the second session has been revealed. Additional details on the experiment, including instructions, can be found in the supplementary online materials.

4.3 Baseline summary statistics

This subsection presents baseline data—first at the level of the neighborhood networks, then at individual level, concluding with balance tests of randomization.

Appendix figure A7 depicts the social network graphs for all neighborhoods, displaying connections between people as directed links. A link is defined when an individual affirms the question "Do you know [$name\ of\ j$]?". The number of isolated respondents who do not know anyone from their neighborhood varies considerably—from eight in figure A7, network (b) to zero in a few others.[31] These graphs give a first 'flavour' of how dense these social networks are. Separated sub-clusters (components) of groups of linked individuals are rare. Most isolated components consist of single nodes without connections to the main component.

Appendix figure A8 displays a different type of network, this time only linking individuals that have exchanged job information or referrals with one another. These job networks are less dense than the networks of people knowing each other. In particular, the number of isolated individuals—i.e. people who have never received job information from or given job information to anyone in their network—is much larger. Generally, the density of the job information exchange seems to closely track the density of the general networks—with neighborhoods where most individuals know each other also displaying a larger number of job information dyads.

Network-level summary statistics

The first panel of table C1 displays network-level summary statistics calculated over all 16 neighborhood networks. The number of individuals per network varies

[31] Note that these network graphs define links if one person claims knowing the other, which is different from requiring that both sides confirm a link—in which case the share of isolated individuals would increase. Unilateral link reporting is a well-known phenomenon in network data (Comola and Fafchamps, 2014).

between 27 and 61, with an average of 46. The average number of general links (two individuals knowing each other by name) per network is quite large at 355, resulting in an average degree centrality of 9.31. This means that on average across all networks, every individual is *directly* connected to over nine other people.[32] There is large variation in the total number of arcs across networks, pointing to substantial differences in how 'dense' the networks are. Indeed, the network density varies strongly from 0.05 to 0.43. The average network transitivity—the chance that two individuals in a neighborhood share a common friend—is very large at 0.61 (61%). In terms of reciprocity r (i.e. the ratio of the number of links pointing in both directions to the total number of links), it is clear that the network is far from being purely bidirectional ($r = 1$), with an average of 0.27, and never exceeding 0.43. Taken together, the average statistics for reciprocity and transitivity suggest that the networks are very dense, but not always covered by bi-directional connections.[33]

The second panel of table C1 displays the same summary statistics for the job information networks, where two individuals are defined as linked when they have exchanged jobs or job information in the past. In line with appendix figure A8, the job exchange networks are less dense than the broadly defined social networks, with lower levels of degree centralisation. The number of components per network is much larger than in the general case, reflecting the larger share of individuals who are isolates or only in a dyadic job exchange relationship with another person. Appendix table C2 presents the network characteristics of table C1 separately for each of the 16 networks, giving an idea of the substantial variation across networks.

Social interactions within neighborhood networks

Moving from network-level statistics to the individual-level, table 2 presents information on the density of social interactions within the neighborhood networks. The average number of links that each person has is approximately 8.7,[34] of which

[32] This is a lower bound: qualitative post-experiment interviews revealed that a substantial share of individuals ($> 50\%$) knows each other by face only, without knowing name or nickname of the other individual.

[33] I can overcome any concerns about the directionality of links by defining a link as undirected, i.e. it is established if either i knows j or j knows i.

[34] This number is different from the network-level summary statistics presented before, because those are not weighted by the number of individuals per network.

she considers three-quarters as relatively central (i.e. as someone to approach if one wanted to spread information about an event in the whole neighborhood, following Banerjee et al. (2019)). On average, individuals spend a total of 227 hours per month with links in their local network, which leads to a substantial average of eight hours per day.[35] Individuals travel to the city centre with approximately a quarter of their connections. People share job information with approximately 30% of their connections, supporting the notion that talking about jobs is an important part of these neighborhood networks. In fact, of those individuals who ever asked another person for help getting a job at baseline, 95% asked directly in their neighborhood, and in 96% of those cases also received meaningful information on jobs. In 40% of these cases, the exchange of job information led to the provision of actual jobs through a direct referral. This demonstrates that actual work, not just information about it, is often obtained through neighborhood networks. The exchange of money is less prevalent and takes place for fewer than 20% of connections.[36] In appendix table C3, I compare these social interactions to the total size of the network. The average respondent knows about 17% of individuals in the network, with substantial variation. At maximum, the number of individuals known directly by a respondent, i.e. connections of degree one, reaches 96% of the whole network.

Who are the central individuals in a neighborhood network? Appendix figure A2 shows standardised coefficients from a multivariate regression of local network centrality on individual characteristics. It shows that being male, part of an ethnic majority in the neighborhood and months lived in the neighborhood are the strongest predictors of centrality, whereas the geographic position in the neighborhood (as measured by distance of the resident's home to the geographic neighborhood centre) does not correlate with network centrality.

[35]The maximum number of hours spent with individuals from the network exceeds the total number of hours per month, because individual i can spend time in groups with several connections j, which would then be summed up separately.

[36]There are small differences in the bi-directional variables for visiting someone's home and job/money exchanges. As the networks are complete and symmetric, the values should in principle be identical. For instance, for every person i lending money to j, j should indicate receiving money from i. Yet, in all cases the values are tilted slightly towards the person visiting or giving the money/information, suggesting that respondents are more likely to remember an active role in an interaction.

Table 2: Connections within the neighborhood networks

Number of links within network:	mean	min	max	sd	count
All links	8.74	0.00	51.00	8.24	739
Considered information spreaders	6.31	0.00	366.00	17.21	739
Total hours spent with links in network	226.85	0.00	3600.00	371.99	739
Travel connections	2.25	0.00	21.00	3.36	739
Links visited	3.51	0.00	31.00	4.96	739
Visitors	3.47	0.00	30.00	4.88	739
Job info given to	2.40	0.00	28.00	3.81	739
Job info received from	2.28	0.00	31.00	3.73	739
Money lent to	1.50	0.00	18.00	2.46	739
Money borrowed from	1.43	0.00	18.00	2.33	739

Notes: Table 2 summarizes social interactions within each of the 16 neighbourhoods, summing all connections of a certain type of individual i in her network.

Individual-level data and balance

Lastly, I present individual-level descriptives. Table D4 shows the balance over the referral treatment groups during the first session of the work experiment. Almost all characteristics are balanced, indicating successful randomisation. In appendix table D5, I compare the individuals participating at the work sessions with those who were surveyed, but never participated. Clearly, people who ever worked in the experiment are different from those who did not—either because the latter did not show up for the first session or because they were not referred by anyone later on. Most importantly, never-workers are more likely to be in (non-permanent) work, less likely to search for (additional) work, and are less connected to other individuals in the sample. In the whole baseline sample, 25% of respondents have worked in the last seven days, with more than 80% having some kind of paid work experience (the rest are first-time job seekers). Approximately two-thirds of the individuals live with their parents and finished school on average three years ago.

5 Empirical strategy

I run two types of regressions to test the different propositions of my model. The first type are cross-sectional regressions, either run separately for each work session or pooled together. Exploiting the random variation in the type of referral treatment assignment, I describe overall trends in the referral strategies and test propositions

1, 3 and 4. The second type of regressions are dyadic regressions looking at direct relationships between every pair of individuals in each network. In the following, I present a typical regression specification for both cross-sectional or dyadic analysis, while specific deviations from these regressions will be introduced and discussed together with the results in section 6.[37]

5.1 Cross-sectional analysis

I run cross-sectional OLS regressions as displayed in specification (7) to test i) whether the performance incentive makes workers refer more productive individuals to the job (proposition 1) and ii) whether the degree of anonymity of the referral leads to the selection of less central individuals (proposition 3 and 4):

$$y_i = \rho_0 + \rho_1 * T_j^k + \theta X_{ij} + u_i \tag{7}$$

where y_i is the outcome of interest of the referred worker i (productivity or network centrality) and T_j^k is the vector of the k different referral treatments of worker j. X_{ij} is a matrix of controls, including session fixed effects and workers' characteristics.

5.2 Dyadic regressions

I run dyadic least squares regression to test the implications of proposition 2, i.e. fewer reciprocal referrals under the incentivized referral treatment condition:

$$\begin{aligned} R_{ij2} = &\beta_0 + \beta_1 R_{ji1} + \beta_2 * Rf_{j1}^{random} + \beta 3 * Rf_{j1}^{incentivized} \\ &+ \beta 4 * Rf_{j1}^{incentivized} * R_{ji1} + \theta X_{ij} + u_{ij}, \end{aligned} \tag{8}$$

where R_{ij2} is an indicator for whether individual i invited person j in session two (and R_{ji1} for whether individual j invited person i in session one). Rf means referral. $\beta_1 > \beta_2$ tests whether i is more likely to refer back to j than a randomly selected member from the neighborhood (reciprocity). $\beta_4 = 0$ tests if the prevalence of reciprocity is the same under incentives (proposition 2 suggests it is lower). All dyadic regressions report standard errors clustered at the dyad level, i.e. at level ij.

[37] I have registered this experiment together with the primary outcomes and analysis of interest in the American Economic Association Registry for randomized control trials under trial number AEARCTR-0002334.

5.3 Outcomes

I measure the effects of the referral type on three main outcomes of interest: i) worker i's productivity, expressed as correctly entered work teams, P_i, ii) worker i's network centrality, measured as the number of (in or out) links at baseline, and iii) reciprocity: an indicator that equals one if worker j refers i to the job in period $t-1$ (i.e. $R_{ji,t-1} = 1$), and i refers j in return in period t (i.e. $R_{ij,t} = 1$).

Measures of network centrality: Different network centrality measures have different characteristics, but all of them try to identify the 'most important' node in a social network. The analysis focusses on degree centrality, i.e. the number of links an individual has. In general social networks, a link is defined as person i knowing another person j by name. I mostly focus on directed networks, i.e. a link between two individuals i and j is defined either by i claiming to know j, $i \rightarrow j$, or vice versa: $j \rightarrow i$. For job information / referral networks, a link is defined as person i having exchanged job information or referrals with another person j in the past, again defined as out-degree ($i \rightarrow j$) or in-degree ($j \rightarrow i$).

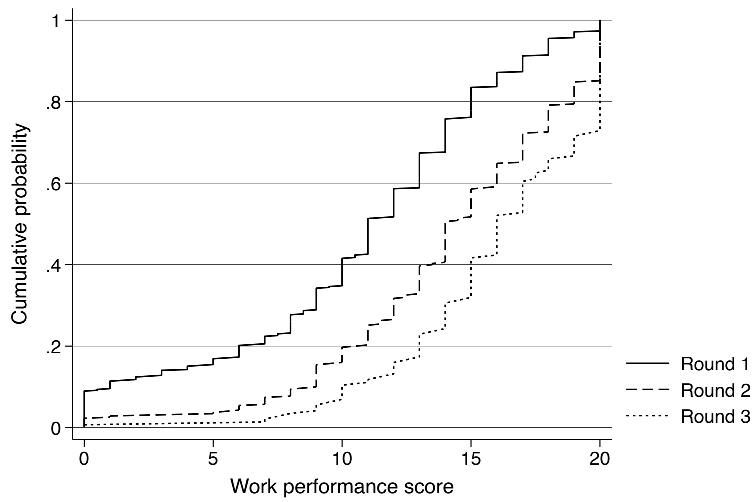
6 Results

This section presents the results of the work experiment. First, I show the data on the open control referral (C). Second, I evaluate the impact of the incentivized referral treatment (T1), testing whether incentives lead to higher worker productivity (proposition 1). In the same step, I introduce the dynamic structure of the experiment to test the effect of reciprocal referring on productivity (proposition 2). Third, I test propositions 3 and 4 by exploiting the variation in the visibility of the referral. In a final step, I look at treatment effects on the network structure, focussing particularly on how isolated workers strategically use the referral opportunity.

6.1 The control condition: open, unincentivized referrals

Figure 1 displays the cumulative distribution functions (CDF) of the workers' performance over the work sessions: the rounds indicate the performance of all individuals working for the first, second, or third time, respectively. With every additional work session workers get better at completing the task without errors.

Figure 1: Cumulative density function of work task performance over time



Notes: This figure presents the CDFs of the individual work task performance (scored on a scale from 0-20) over subsequent work sessions.

The median payment (calculated as performance score $\times 10$) for an individual's first work stint is 120 ETB, and increases to 150 ETB (170 ETB) in stint two (stint three). The CDFs for repeatedly participating in the work activity stochastically dominate the previous day's CDF. I can reject equality of the distributions with a two-sample Kolmogorov-Smirnov test (for all three comparisons $p < 0.01$). The improvement in work performance shown in this graph can be explained by two different factors: on-the-job learning and (incentivized or unincentivized) selection into the work. The extent of within-worker improvement through 'learning on the job' needs to be taken into account when looking at the impact of referral type on job performance, as some workers will have already worked in a previous session.

6.2 Variation in the monetary pay-off of the referral (T1)

I test whether an increase in the incentive payment P_i leads to a selection of better-performing individuals (proposition 1), with a regression similar to equation (7).[38] Due to the piece-rate payments, all workers face private incentives to do well at work, regardless of the 'referral contract'. Since the referral treatment status

[38] Specifically: $P_i = \rho_0 + \rho_1 * T_j^{incentivized} + \rho_2 * T_j^{random} + \theta X_{ij} + u_i$.

(incentivized vs non-incentivized referral) is randomly allocated to the previous session's workers, all differences in their referrals performance can be attributed to the inviting worker's extra incentive to invite particularly well-performing people.[39]

I find that incentivising the previous session's worker to bring in a more productive individual to the work sessions leads to an improved performance of the referral (table 3). This results is in line with Beaman and Magruder (2012)'s finding. Referrals invited by an incentivized worker score approximately 2.4 points higher on the performance schedule than referrals invited by workers without incentives. In relative terms, this means an improvement of approximately 18% compared to the control referral performance (C). This finding holds across different specifications.[40]

Since some workers already worked in the first round (and the results are based on performances in the subsequent sessions), the positive effect on incentivized referrals could partly be driven by on-the-job learning, if the incentivized referral disproportionately leads to the referral of prior workers (or if learning is asymmetric between the workers invited under different treatments). Though there is no statistical difference in whether the invited worker had already participated previously by referral incentives, I formally separate the learning effect from the ability effect by imputing the workers first-round productivity, i.e. the performance on the job when they worked the first time. Column (4) of table 3 shows that the effect of the incentivized referral is even larger, indicating that on-the-job learning if anything works against pure ability: on-the-job learning of workers invited without incentives decreases the productivity gap by a third over time.

These findings demonstrate that workers can be properly incentivized to screen

[39]While Beaman and Magruder (2012) look at low-stakes and high-stakes incentive payments, I only have one payment category. Since the performance payments of the original participant in my work sessions is matched to be 100% of the referral's performance payment, this can be characterized as "high-stakes". Beaman and Magruder (2012) also do not pay their participants according to performance, which leads to smaller incentives to perform well during the task. Here, workers are paid by their performance, which arguably is closer to real-life day-work markets.

[40]I can rule out that workers who were incentivized in previous rounds tell their referrals about how to do the work best, since a) the work materials are kept in the session rooms, and b) the content of the task varies over the rounds.

Table 3: The impact of incentivized referrals on work performance

	Work performance	Work performance	Work performance	Work performance net of learning
Incentivized referral$_{ji}$	2.354***	2.385***	2.385***	4.435**
	(0.853)	(0.836)	(0.838)	(2.167)
Constant	15.26***	14.30***	15.44***	11.08***
	(0.671)	(0.682)	(0.931)	(1.039)
Invited by computer$_i$	Yes	Yes	Yes	Yes
NH fixed effects	Yes	Yes	No	No
Day fixed effects	No	Yes	No	No
Session fixed effects	No	No	Yes	Yes
Number of observations	664	664	664	491
p-val. for H0: Incent. referral$_{ji}$=Invited by comp.$_i$	0.001	0.000	0.001	0.015

Notes: Table 3 presents a regression of a worker i's performance on the type of referral contract (incentivized, by computer or non-incentivized control). This table pools worker i's performances from sessions two and three. Session fixed effects implicate neighbourhood and day fixed effects. Column 4 imputes worker i's's score with her score from the first time she completed the task. Standard errors in parentheses. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level.

more productive individuals for a job — both in general (the main result in Beaman and Magruder, 2012), and when doing so in their local network. This is the case even when their referral's performance is independently incentivized. This finding presents a higher threshold than screening more productive individuals for an un-incentivized work task, and minimised the impact of side payments among worker and referral, as referrals have sufficient private incentives to perform well.

6.3 Exploiting the referral panel: repeated interactions

I now exploit the multiple rounds of the experiment to look at dynamic referral contracts. In particular, the dynamic structure of the experiment allows me to look at one specific component of the social payoff σ_{ij}, namely reciprocal referrals ρ_{ji}. Proposition 2 states that an increase in the performance pay P_i should lead to lower social payoffs σ_{ij}. Therefore, one can expect to see fewer reciprocal referrals in the incentivized referral treatment condition.

Following specification (8), I regress an indicator variable for whether worker i invited person j after the second work day (i.e. to the third work session) on whether j had previously referred i to the work. Table 4 presents the results. The first column is a reduced version of specification 8, which is run on the partially selected sample of workers present in the second session. I find that a prior work referral from j to i is a very strong predictor of whether subsequently i re-invites j

Table 4: Dyadic regressions: Individuals refer strongly reciprocally

	(1)	(2)			(3)	
	Referral$_{ij2}$	Referral$_{ij2}$	Referral$_{ij2}$	Referral$_{ij2}$	Referral$_{ij2}$	Referral$_{ij2}$
Referral$_{ji1}$	0.286***	0.287***	0.280***	0.279***	0.279***	0.288***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Random$_{i1}$		0.00469**	0.00507**	0.0104***	0.0104***	0.0105***
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Incentivised$_{i1}$					0.00141	0.00325
					(0.00)	(0.00)
Incentivised$_{ji1}$						-0.0644***
						(0.02)
Baseline controls	No	No	Yes	Yes	Yes	Yes
NH fixed effects	No	No	No	Yes	Yes	Yes
Number of observations	17280	17280	17280	17280	17280	17280
p-value for H0: Referral$_{ji1}$=Random$_{i1}$		0.000	0.000	0.000	0.000	0.000
p-value for H0: Referral$_{ji1}$=Incentivised Referral$_{ji1}$						0.000

Notes: Table 4 presents the results from specification (8), regressing an indicator variable for whether worker i invited person j after the second work day on whether j had previously referred i to the work. Standard errors in parentheses. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level.

— it increases the chances of the latter happening by 29 percentage points. In order to causally interpret these coefficients, I compare this coefficient to the chances of being invited by a randomly selected neighborhood worker i.[41] The middle panel (columns 2 to 4) of table 4 shows that after including the dummies for being referred by a random worker, the results on the reciprocity indicator Referral$^1_{ji}$ remain unchanged. The coefficient comparison of Referral$^1_{ji}$ to the likelihood of being invited by a random individual is statistically highly significant throughout. This suggests that, compared to a random worker, workers are much more likely to refer a person to the job who has previously invited them. This finding points toward strong reciprocity at the work place in this context—in contrast to Conte et al. (2015) who find that only a minority of players choose a reciprocal referral strategy in lab-referral games played with European students. In the local labor markets of this paper, where individuals from real social networks provide each other with work opportunities, reciprocity in job referrals accounts for a large share of the strategic network interactions.

The third panel of regressions of table 4 shows that the prevalence of reciprocal referring is driven by non-incentivized referrals. For incentivized referrals, the effect on reciprocity is still positive, but smaller than for the un-incentivized treat-

[41] A subsample of workers is randomly invited to the second work session.

ment. In particular, when interacting the round referral$^1_{ji}$ with whether j was in the incentivized treatment group, the effect on reciprocity is about a fifth smaller than for the case of non-incentivized treatments. This is evidence in line with proposition 2: while a round one referral from j to i is still an important predictor of a round two referral from i to j, the prevalence of such direct reciprocity is 20% lower under incentives. This supports the idea that an increase in the performance payment P_i leads to a decrease in reciprocal referrals ρ_{ji} within the social payoff σ_{ij}.

Reciprocity and the implications for productivity

Is the strong prevalence of reciprocal referrals detrimental for productivity? A related literature shows that under referral schemes, in-group biases and ethnic differences can prevent an efficient allocation of workers to jobs (Alesina and Ferrara, 2005, Anderson, 2011, Hjort, 2014, Burgess et al., 2015, Fisman et al., 2017). Similarly, in this experiment, a large share of the reciprocity is driven by co-ethnic workers referring each other, raising the question of how this affects output. Conceptually, proposition 2 demonstrates the trade-off between social payoff σ_{ij} and the selection of referees. An increase in P_i leads to a selection of more productive workers as well as a decrease in the prevalence of reciprocal referral arrangements.

I test this relationship between reciprocity and productivity by regressing work performance scores of the third (and last) work session on whether the worker i has been invited reciprocally, randomly (i.e. by the computer), or differently (e.g. by another worker, but not in a reciprocal fashion). Table 5 shows the results. I find that compared to non-reciprocally invited workers, both randomly and reciprocally invited workers perform worse. Since "non-reciprocally" invited workers might be selected positively on ability, the true comparison of interest here is whether reciprocally invited workers perform worse than randomly invited ones.[42] Across specifications, I find that reciprocal referring comes at a substantial productivity cost.

[42]Reciprocally invited workers have, by definition, worked in at least one previous session and are thus familiar with the task. This could mean better performance through learning on the job (figure 1). Alternatively, workers who do a similar job twice may get bored and perform worse, despite performance pay. Regardless I control for both directions of repeated work by imputing the first-session performance of all workers in the sample, which directly captures individual work ability.

On average, reciprocally invited workers perform 0.5 points worse than randomly invited ones, which is a productivity decrease of 5% over a baseline control mean of 10.4. The difference between reciprocal and random worker performance is statistically significant across specifications. Compared to non-random, non-reciprocally referred workers, which is a more natural comparison to most labor markets where workers are not invited randomly, the decrease in performance is even more substantial, with a decrease of 1.25 points (11%) over a control mean of 11. Importantly, the result that reciprocal links are less productive still holds when controlling for whether a referral was incentivized (Table 5, last column). While the difference between reciprocally and not-reciprocally invited individuals under incentivized referrals is also negative and significant, the performance *levels* of workers invited under incentives are 2.5-4 points higher. Interpreted through the lens of my framework the immediate reciprocation of referrals, ρ_{ji}, is the ingredient of the social payoff σ_{ij} that makes individuals select lower-performing individuals compared to the direct monetary incentive payoff. Overall, the findings suggests that reciprocal job referring leads to important productivity losses in these local labor markets, as the reciprocally referred worker often is not the most productive one.

6.4 Variation in the social pay-off of the referral (T2 & T3)

I now turn to the impact of changing the visibility of the job referrals induced by treatments 2 and 3. In a first step, I describe the correlation between measures of network centrality and the likelihood of being referred to the job under the open referral (C). Second, I test how varying the importance of κ_{ji} and ρ_{ji} in the network benefits function impacts the centrality c_j of the referral (propositions 3 and 4).

6.4.1 Open referrals: the role of network centrality

To investigate the role of network centrality, I regress a binary indicator of whether individual i referred individual j to the job in open referrals (C) on the baseline number of connections of j, her degree centrality in her neighborhood network, and on a variety of further network centrality measures. The regressions are run on the cross-sectional data (i.e. one observation per individual). Importantly, the coefficients in these regressions cannot be interpreted in a causal sense, since network centrality correlates with a host of other characteristics of individuals.

Table 5: Dyadic regressions: reciprocal referring and costs for productivity

	Work performance score$_i$			
Reciprocally invited$_i$	-1.265***	-1.209***	-1.154***	-1.114***
	(0.13)	(0.13)	(0.13)	(0.13)
Randomly invited$_i$	-0.795***	-0.655***	-0.767***	-0.758***
	(0.11)	(0.11)	(0.13)	(0.13)
Incentivised and reciprocally invited$_i$				-1.763**
				(0.77)
Incentivised invited$_i$				4.085***
				(0.57)
Constant	12.62***	14.94***	17.58***	17.49***
	(0.06)	(0.37)	(0.45)	(0.45)
Baseline controls	No	Yes	Yes	Yes
NH fixed effects	No	No	Yes	Yes
Number of observations	9990	9990	9990	9990
p-value for H0: Score$_{i,reciproc.}$=Score$_{i,random}$	0.001	0.000	0.021	0.034
p-value for H0: Score$_{i,reciproc.,incentiv.}$=Score$_{i,reciproc.}$				0.000

Notes: Table 5 displays the results of an OLS regression of work performance in the third (and last) work session on whether worker i has been invited reciprocally, randomly (i.e. by the computer), or differently (e.g. by another worker, but not in a reciprocal fashion). The work scores display the first-session performance of the workers, which directly captures individual work ability rather than learning on the job. Standard errors in parentheses. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level.

The results are shown the top panel of appendix table E8. When the different centrality measures enter the regression separately, both out-degree and in-degree centrality are significant predictors of whether an individual receives a referral to a job, but the coefficient for in-degree centrality is 20% larger. An increase of j's in-degree centrality by one person (i.e. being known by one more individual in the neighborhood) would increase her chances of being referred to the job by 1.2 percentage points, or approximately 4%. When including both in- and out-degree as regressors, the in-degree centrality remains significant and more than twice as large as the out-degree. In the regression that includes all centrality measures simultaneously (last column), a similar picture emerges: in-degree centrality is the relevant measure predicting the chances of being referred to a job, i.e. the number of individuals claiming to know you. This makes intuitive sense: since the individuals j are referred to the job (by individuals i), it matters how many people the referral j is known by—which is j's in-degree centrality. It is less important how many other people in the neighborhood j claims to know—j's out-degree centrality—because the inviter i is making the referral. I repeat the same analysis for job network centrality in the bottom panel of appendix table E8. The findings are very similar to those for general networks and discussed in appendix section E.6.

6.4.2 Why do workers refer central nodes? Varying the referral visibility

A worker's in-degree centrality positively correlates with her likelihood of being referred to the job. At the same time, I find no evidence that central workers perform better on the job, even under incentivized referrals (see appendix section E.2). This rules out that inviting a central person to the job increases the direct monetary pay-off for the referee through the incentive treatment or side payments. In this subsection, I present three pieces of evidence that more central workers are more likely to get job referrals in their networks due to the importance of potential future network benefits κ_{ji} compared to immediate reciprocity concerns ρ_{ji}. First, I demonstrate that under partial anonymity, workers invite less central individuals compared to the open referrals. Second, I show that this tendency increases when all potential network benefits are switched off, under full anonymity. Lastly, I present short-run follow-up data collected a few days after the experiment, where workers state why they made a given referral. I show that self-stated other-regarding preferences increase with the amount of anonymity in the treatment.

In a first step, I test if an increase in the importance of return referrals ρ_{ij} compared to future benefits κ_{ji} will lead to the selection of workers with lower centrality (proposition 3). To do this, I compare the referrals workers make in their own neighborhood to those made partly or fully anonymously to other neighborhoods, where I control which information are made available to the participant about personal characteristics of the unknown referrals.[43] The partly anonymous referral treatment (T2) in other neighborhoods explicitly encourages reciprocity ρ_{ij}, but is unlikely to generate any lasting future exchanges κ_{ji}. In a second step, I test if when network and monetary benefits are ruled out, workers will select referrals with the lowest centrality (proposition 4). The fully anonymous referral treatment

[43] These treatments T2 and T3 are based on within-subject variation. The partly anonymous referral encourages the referred worker to reciprocate. In the anonymous referral, the invited workers does not know anything about who referred her.

(T3) in other neighborhoods completely rules out reciprocity or further benefits ($\rho_{ij} = \kappa_{ji} = 0$), so that the social payoff reduces to the inclusion of j's income in i's social payoff: $\sigma_{ij} = \omega_i m_j$. Both tests are conducted with the following regression:

$$Centr_j = \alpha_0 + \alpha_1 \cdot Rf_i^{own} + \alpha_2 \cdot Rf_i^{other,p.anon.} + \alpha_3 \cdot Rf_i^{other,anon} + \theta X_j + u_{ij}, \quad (9)$$

where Rf_i^{own} is a dummy for the referral in i's own neighborhood, while $Rf_i^{other,p.anon.}$ and $Rf_i^{other,anon}$ indicate partly anonymous and fully anonymous referrals in other neighborhoods. Propositions 3 and 4 suggest that $\alpha_1 > \alpha_2 > \alpha_3$.[44]

Table 6 displays the results. Comparing the referrals that workers make to people in their own social network to the referrals these same workers make to (unknown) individuals from other, similar neighborhoods, I find contrasting patterns. In the own neighborhood networks, referred individuals are positively selected on centrality, whereas the opposite is the case in other neighborhoods. The difference between partly anonymous referrals in other neighborhoods (enabling reciprocity) and anonymous referrals in other neighborhoods is negative and statistically significant. Why do workers invite less central individuals when referring out of their neighborhood, but more central individuals in their own network? The framework suggests that *ceteris paribus* reciprocity is more likely from individuals with fewer connections (T2), especially when other network benefits can be ruled out. In the case of fully anonymous referrals (T3), network and monetary benefits can be ruled out, so that only other-regarding preferences remain. It follows that, in line with proposition 4, individuals with the fewest connections are referred.

To conclude this section, I present descriptive evidence on why workers refer more central individuals to the job, drawing from follow-up data collected a few days after the experiment. Specifically, I characterize the expectations workers have about central individuals' earnings and job information, and I analyse the self-reported reasons why workers made a specific referral. Starting with the expectations, the left panel of figure 2 shows that workers overwhelmingly expect individuals in their neighborhood who have more connections to earn more per

[44] The comparison of α_1, α_2, and α_3 is based on experimental variation in referral type, whereas the comparison of each coefficient to the omitted category (non-selected workers) is based on selection via referral.

Table 6: Centrality of referred workers compared to whole network

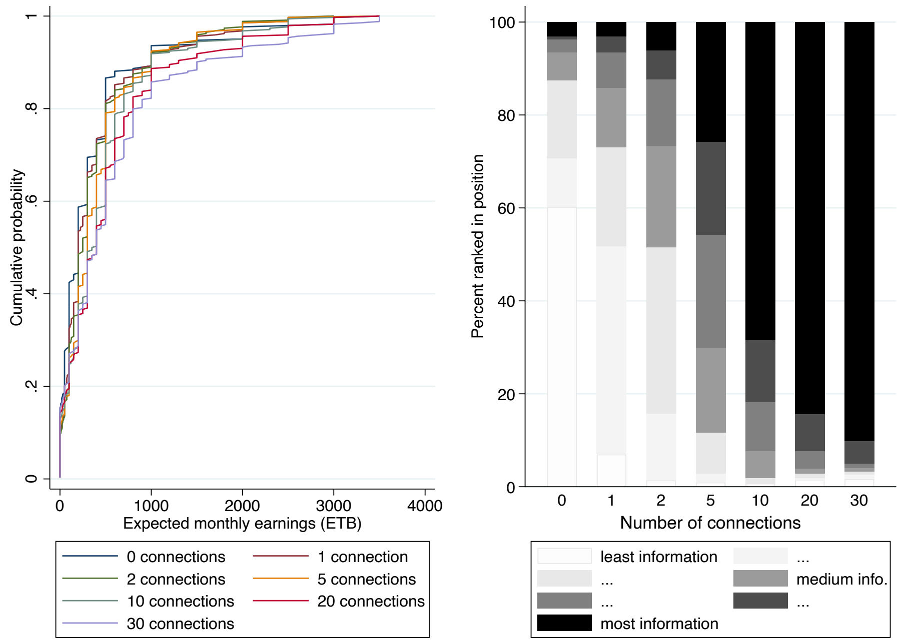
	Number of connections of invited worker j (S.E.)
Referral in own neighbourhood (C)	2.644*** (0.88)
Open referral in other neighbourhood (T2)	-3.742*** (0.68)
Anonymous referral in other neighbourhood (T3)	-4.444*** (0.62)
Mean number of friends (SD)	10.53 (8.95)
Number of observations	991
p-value for H0: Open referral in own neighbourhood= Open referral in other neighbourhood	0.000
p-value for H0: Open referral in other neighbourhood= Anonymous referral in other neighbourhood	0.039

Notes: Table 6 shows the results of specification (9), an OLS regression of invited worker j's degree centrality on the (exogenous) type of referral treatment of the inviting worker i. Data from all work sessions is pooled. Treatments T2 and T3 are based on within-subject variation. The comparison of treatments C, T2 and T3 to each other is based on experimental variation in referral type, whereas each coefficient is based on selection via referral. Standard errors in parentheses. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level.

month. The CDFs for higher number of connections monotonically dominate the lower ones. The right panel of figure 2 shows that the respondents expect individuals from their neighborhood with more connections to also have access to more information about job opportunities. While the increase is monotonic, it is not linear, with large jumps in the information ranking between five and ten connections.

Next, appendix figure A6 displays the self-reported reasons why workers made a referral to a specific person. The left panel includes more general answer categories. In line with my experimental results, workers are more likely to claim they referred a more central person to the job in the open, own-neighborhood conditions, compared to the referrals made to individuals from other neighborhoods. For the latter two, the self-reported reasons do not differ significantly by the type of referral. Both in the partly and fully anonymous referral type, workers claim to have been guided a lot by the other person's gender or age. In the own neighborhood, many respondents claim to have been mostly guided by whether another workers was a baseline connection. The right panel of figure A6 zooms in on the answers that specifically

Figure 2: Workers' expectations about correlates of centrality



Notes: This figure presents individual-level expectations about the correlates of local network centrality. The expectations are fixed about 'someone living in your neighborhood'. The left panel shows the CDFs of how much an individual is expected to earn each month, by how many neighborhood connections this person has. The right panel shows the ranking of how much labor market information an individual is expected to have, again by number of neighborhood connections.

mention the referral's centrality. Of those people who claim to have referred a more central person, most mention the job information central people have as the main reason, followed by the idea that central people would make easy social connections. Among workers stating they referred a less central individual, more than half base this on pity—in particular in the other neighborhood treatments—followed by the idea that it would be easy to connect to those individuals.

Taken together, these self-reported follow-up findings are in line with the conceptual framework: connecting to central people in the own network dominates connecting to peripheral nodes, due to the long-run information benefits (κ_{ji}) associated with central workers, whereas in other networks immediate reciprocity concerns (ρ_{ji}) or other-regarding preferences ($\omega_i m_j$) dominate.

7 Discussion

In this section, I discuss the implications of my previous findings on equality in local labor markets. I have shown in the previous section that open, un-incentivized referrals—the ones most often encountered in the field—lead to the selection of more central individuals to jobs, and are characterized by reciprocal invitations. Both of these factors prevent peripheral individuals in the local network from accessing jobs. This begs the question whether and how such individuals, without many connections, can be included in the local labor market. I provide two pieces of evidence describing how peripheral workers can or cannot get access to referral networks. First, I demonstrate that peripheral individuals are more likely to form new relationships through the work experiment. And second, I show that these new referral links are more likely to be reciprocated by the other worker (compared to connections that already existed at baseline) and that a substantial proportion of new connections are still in place one and a half years after the experiment.

7.1 How persistent is exclusion?

The previous findings show a strong positive correlation between individual network centrality and the likelihood of being referred to the job. The effect is driven by an individual's in-degree centrality, i.e. by how many other people she is known, not by how many individuals she knows. This suggests that the long-run benefits of being connected to person j, κ_{ji}, dominate compared to the short-run reciprocity concerns ρ_{ji}. At the same time, reciprocal referrals are very prevalent in the labor market even for central individuals. Both findings both point to a disadvantage for individuals who are not central in their local network and who do not have the opportunity to provide jobs or favours themselves.[45] Given the baseline network structure (figure A7), individuals without many connections in their local job network face a challenging situation—in particular, when in-degree centrality is the driver behind referrals, as this is something individuals are less likely to be aware of and in control of than out-degree centrality (which is reported by themselves).

In this step, I test if peripheral individuals can use the intervention to gener-

[45]Controlling for whether two individuals know each other at baseline shows that this is a very strong predictor of a referral from one worker to another.

ate new links when (randomly) given the opportunity to distribute jobs in their network. I run the following regression:

$$R_{ij}^{new} = \zeta_0 + \zeta_1 Rf_i + \zeta_2 Rf_i \cdot PERI_i + \theta X_{ij} + u_{ij} \quad (10)$$

where R_{ij}^{new} is an indicator variable for a newly established link (via referral) between individuals i and j (who were thus not linked at baseline). $Rf_i \cdot PERI_i$ is an interaction between the exogenous referral treatment and whether the individual i is at the periphery of her local network (with varying definitions). This regression is run on the pooled data from different rounds. The test $\zeta_1 = \zeta_2$ shows whether peripheral individuals i are more likely to establish new links with other people from their neighborhood than non-peripheral people. Table 7 presents the results.

I find that treated individuals (i.e. those making a referral in the first work session) from the periphery (here defined as having < 2 baseline connections) are significantly more likely to establish a connection to individuals not known at baseline, by approximately 1.5 percentage points. The coefficient is small, but represents a sizeable increase compared to non-peripheral treated workers. Establishing new connections is rare, but it happens more than twice as often for people at the periphery of the networks. Given that approaching unknown individuals and presenting them with an invitation to a work opportunity might involve high barriers, the coefficient should be interpreted as a lower bound of how peripheral individuals are more likely to make new connections rather than relying on their existing ones.

The right hand side of table 7 displays results for the job information networks, where two people are linked if they shared jobs or job information in the past. As a consequence, the share of peripheral individuals increases considerably (cf. figure A8). Regardless, I still find a significant effect of 1.2 percentage points for treated individuals, which is roughly a doubling of the coefficient for treated, but non-peripheral individuals. Taken together, the results indicate that individuals with few or no connections in the general or the job exchange networks utilize the referral opportunity to connect with new people. They use the 'windfall' invitation to a real and well-paid day job for a strategic expansion of their social network.[46]

[46] In appendix figure A4, I show that both previous results are not driven by a mechanical relationship of how I define periphery.

Table 7: Dyadic regressions: persistent exclusion from general and job networks

	General networks			Job networks		
	New Link$^1_{ij}$	New Link$^1_{ij}$	New Link$^1_{ij}$	New Link$^1_{ij}$	New Link$^1_{ij}$	New Link$^1_{ij}$
Treatment$_i$	0.0101***	0.0108***	0.0100***	0.0128***	0.0136***	0.0133***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Peripheral$_i$*Treatment$_i$	0.0174***	0.0165***	0.0142***	0.0127***	0.0121***	0.0112***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Baseline controls	No	Yes	Yes	No	Yes	Yes
NH fixed effects	No	No	Yes	No	No	Yes
N	17280	17280	17280	17280	17280	17280

Notes: Table 7 presents the results of specification (10), where I regress an indicator variable for a newly established link (via referral) between individuals i and j on the treatment status as well as an interaction between the exogenous referral treatment and whether the individual i is at the periphery of her local network. Here, i is defined as peripheral when having fewer than two out-degree connections at baseline. The top panel shows the results for general networks, the bottom panel for job networks. Standard errors in parentheses. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level.

Now, with peripheral individuals trying to expand their connections by referring individuals previously not in their network, are they successful in consolidating these new links? That is, once they have provided a new connection with a day job opportunity, do those individuals pay back the favour through reciprocal referring, as previously found for the whole sample (cf. table 4)? Appendix table E9 shows the results for general and job networks. I find that peripheral individuals that invite a new connection to the day job are *more likely to be reciprocated* than non-peripheral individuals. This is both in comparison to a randomly invited worker, but also in comparison to peripheral individuals that did not establish a new link in the previous session (but rather invite one of their few baseline connections). So peripheral peers that make new connections not only manage to consolidate that new link by being re-invited, they are actually better off than similar peripheral nodes that do not refer an unknown individual after the first work session.

7.2 The long-run effects of the day job intervention

In the previous subsection, I have established that workers who are peripheral in their social network are more likely to use the treatment to establish new connections in the neighborhood, and that these new connections also get reciprocated immediately (ρ_{ji}). As a final puzzle, however, it is unclear whether these new connections also manage to generate long-run network benefits for the workers who are peripheral at baseline (κ_{ji}), or whether the new links break after some time.

In order to look at the long-run impacts, I re-interview the respondents 18 months after the work sessions about those links that were newly created in the experiment.[47] The left panel of figure 3 presents the 18 months follow-up data on new connections: of those individuals who made a referral to a previously unknown person and could be reached during the follow-up, 75% still know that person, 42% are still in contact with that person, and 24% are still in a job-exchange relationship with that person, with 24% of respondents giving information and 15% receiving information about jobs. These numbers indicate that a low-touch, short intervention can generate new and lasting relationships between individuals.

Do these persisting relationships simply result from secular changes in connectivity over time? Two pieces of evidence suggest differently: first, the right panel of figure 3 compares workers who made one new connection in the experiment with those who made two. In line with the idea that these new relationships were established *during the experiment* and do not represent secular network change, workers who made two new connections are 1.5 times more likely to still be in contact with at least one of the them, compared to workers who established only one new link. This holds across most interaction types. Second, the baseline data suggests that for every month someone lives in the neighborhood, she makes 0.012 new connections, corresponding linearly to a secular change of 0.216 new connections over 18 months. The long-run follow-up data yield a number at least twice as large (42% still in contact), corroborating the idea that these are new connections generated in the experiment, thus strengthening the position of baseline peripheral workers.[48]

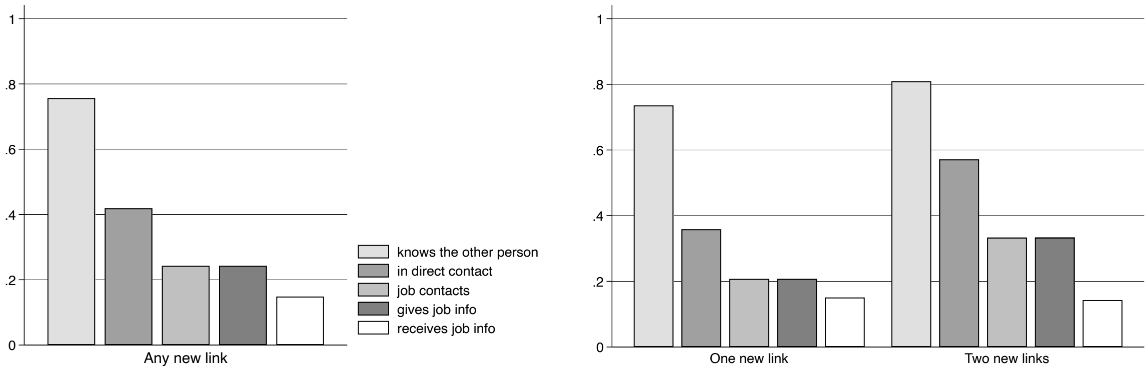
8 Conclusion

This paper studies the motivations behind workers' job referral decisions. To do so, I hire individuals—out of permanent employment and living in densely popu-

[47]I identify 84 individuals who made a referral to a previously unknown person in the experiment. Of those, 75 could be reached, which equals an attrition rate of 10.5% after 18 months. Of those respondents who could not be reached, six migrated out of Ethiopia and the rest did not have any working phone number. One respondent refused participation in the long-run follow-up.

[48]Appendix section E.5 shows robustness to various periphery definitions.

Figure 3: Persistence of newly established social connections after 18 months



(a) General networks (b) General networks, by # of new links

Notes: This figure presents the 18 months follow-up data on respondents who established new connections. The figure shows the share of respondents still engaged in different types of social interactions, pooled (left) or split by whether the original respondents approached one or two new links during the experiment (right).

lated neighborhoods in Addis Ababa, Ethiopia—for day work in a data entry firm. Upon work completion, the workers can invite social contacts to the same job. I vary the conditions of this job referral process along two key dimensions: whether the referral is monetarily incentivized or not, and whether the referral is visible or not. The work is conducted over several work days, with repeated referral opportunities.

This paper makes several contributions. First, I find that the job referrals are characterized by a substantial degree of reciprocity, which is significantly reduced under the incentivized referral treatment. This finding, together with the fact that reciprocally invited workers are less productive, demonstrates the trade-off between direct referral performance incentives and (both short-run and long-run) social payoffs—a trade-off which I derive more formally in a conceptual framework. In line with the model, I find that when given the performance incentive, workers screen more productive individuals in their local neighborhood networks, utilizing information about other network members. The prevalence of reciprocity under the open, non-incentivized referral treatment also implies that individuals with one-off referral opportunities have larger chances of receiving casual work in the future.

In addition, I find that the centrality of the referred workers varies substantially with the type of referral treatment. Under the standard open referral treatment, more central individuals are invited to the job, even though they are not more productive. This again means that individuals without many connections are left jobless. In contrast, under the referral treatments which are partly or fully anonymous, less central workers are selected—due to the reduced importance of network benefits and dominance of other-regarding concerns.

Finally, I show that there is persistence in social network exclusion. Individual connectedness in neighborhood networks predicts the likelihood of obtaining jobs in informal local labor markets, and peripheral individuals who are not (randomly) given a day job and a referral opportunity largely remain excluded. This finding could provide an alternative explanation for long-term unemployment and detachment from the labor market, in particular in the informal and local labor markets of many cities in developing countries. On a more optimistic note, peripheral individuals who are (randomly) given referral opportunities manage to use these as devices to connect to larger components of the neighborhood networks, by establishing links with previously unknown individuals. These links are usually confirmed by the other node through a reciprocal job invitation, and hence remain in place after more than one work session. One and a half years after the experiment, over 40% of new social connections are still intact, demonstrating the potential of a very light-touch intervention to affect connectedness in social networks in the long-run.

This paper suggests that is possible to design markets that enable socially excluded (but strategically sophisticated) individuals to integrate into networks. Individuals can permanently overcome exclusion once given the chance—this finding has implications for how policy makers could think about alleviating unemployment, e.g. through the provision of subsidised temporary job opportunities. Ultimately, I find that a rather brief provision of work and referral opportunities has long-run effects on connectedness in the local labor market. This conclusion goes beyond the context of cities in the developing world and could similarly apply to rich countries, where economic development is often said to have brought about an erosion of social interactions (Putnam, 2000).

References

Abebe, G., S. Caria, M. Fafchamps, P. Falco, S. Franklin, and S. Quinn (2021). Anonymity or Distance? Job Search and Labour Market Exclusion in a Growing African City. *Review of Economic Studies*.

Abebe, G., S. Caria, M. Fafchamps, P. Falco, S. Franklin, S. R. Quinn, and F. J. Shilpi (2020). Matching Frictions and Distorted Beliefs. Working paper.

Abebe, G., S. Caria, and E. Ortiz-Ospina (2020). The Selection of Talent. Mimeo.

Alesina, A. and E. L. Ferrara (2005). Ethnic Diversity and Economic Performance. *Journal of Economic Literature 43*(3), 762–800.

Anderson, S. (2011). Caste as an Impediment to Trade. *American Economic Journal: Applied Economics 3*(1), 239–263.

Bala, V. and S. Goyal (2000). A Noncooperative Model of Network Formation. *Econometrica 68*(5), 1181–1229.

Bandiera, O., I. Barankay, and I. Rasul (2009). Social Connections and Incentives in the Workplace: Evidence From Personnel Data. *ECMA 77*(4), 1047–1094.

Banerjee, A., A. G. Chandrasekhar, E. Duflo, and M. O. Jackson (2019). Using Gossips to Spread Information: Theory and Evidence from Two Randomized Controlled Trials. *The Review of Economic Studies 86*(6), 2453–2490.

Bayer, P., S. L. Ross, and G. Topa (2008). 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 (2018). Do Job Networks Disadvantage Women? Evidence from a Recruitment Experiment in Malawi. *Journal of Labor Economics 36*(1), 121–157.

Beaman, L. and J. Magruder (2012). Who Gets the Job Referral? Evidence from a Social Networks Experiment. *American Economic Review 102*(7), 3574–3593.

Bolton, G. E. and A. Ockenfels (2000, March). ERC: A Theory of Equity, Reciprocity, and Competition. *American Economic Review 90*(1), 166–193.

Breza, E. and A. Chandrasekhar (2019). Soc. Networks, Reputation, and Commitment: Evidence from a Savings Monitors Experiment. *ECMA 87*(1), 175–216.

Burgess, R., R. Jedwab, E. Miguel, A. Morjaria, and G. Padró i Miquel (2015). The Value of Democracy: Evidence from Road Building in Kenya. *American Economic*

Review 105(6), 1817–1851.

Cai, J., A. De Janvry, and E. Sadoulet (2015). Social Networks and the Decision to Insure. *American Economic Journal: Applied Economics 7*(2), 81–108.

Cai, J. and A. Szeidl (2017). Interfirm Relationships and Business Performance. *The Quarterly Journal of Economics*, qjx049.

Calvó-Armengol, A. (2004). Job Contact Networks. *Journal of Economic Theory 115*(1), 191–206.

Calvó-Armengol, A. and M. O. Jackson (2004). The Effects of Social Networks on Employment and Inequality. *American Economic Review 94*(3), 426–454.

Caria, A. S. and I. W. Hassen (2013). The Formation of Job Referral Networks: Experimental Evidence from Urban Ethiopia. Discussion paper 1282, IFPRI.

Caria, S., S. Franklin, and M. Witte (2020). Searching with Friends. IZA DP.

Chandrasekhar, A. G., C. Kinnan, and H. Larreguy (2018). Social networks as contract enforcement: Evidence from a lab experiment in the field. *AEJ: Applied Economics 10*(4), 43–78.

Chandrasekhar, A. G., M. Morten, and A. Peter (2020). Network-Based Hiring: Local benefits; Global Costs. *Working Paper*.

Charness, G. and M. Rabin (2002). Understanding Social Preferences with Simple Tests. *The Quarterly Journal of Economics 117*(3), 817–869.

Chuang, Y. and L. Schechter (2015). Social Networks in Developing Countries. *Annual Review of Resource Economics 7*(1), 451–472.

Comola, M. and M. Fafchamps (2014). Testing Unilateral and Bilateral Link Formation. *The Economic Journal 124*(579), 954–976.

Conte, A., D. T. Di Cagno, and E. Sciubba (2015). Behavioral patterns in social networks. *Economic Inquiry 53*(2), 1331–1349.

Dhillon, A., V. Iversen, and G. Torsvik (2020). Employee Referral, Social Proximity, and Worker Discipline. Working paper.

Dustmann, C., A. Glitz, U. Schönberg, and H. Brücker (2016). Referral-Based Job Search Networks. *The Review of Economic Studies 83*(2), 514–546.

Eom, Y.-H. and H.-H. Jo (2014). Generalized friendship paradox in complex networks: The case of scientific collaboration. *Scientific reports 4*(1), 1–6.

Fafchamps, M. and S. Quinn (2016). Networks and Manufacturing Firms in Africa: Results from a Randomized Field Experiment. *World Bank Economic Review*.

Fehr, E. and S. Gächter (2000). Fairness and Retaliation: The Economics of Reciprocity. *Journal of Economic Perspectives 14*(3), 159–181.

Fehr, E. and K. M. Schmidt (1999). A Theory of Fairness, Competition, and Cooperation. *The Quarterly Journal of Economics 114*(3), 817–868.

Feigenberg, B., E. Field, and R. Pande (2013). The Economic Returns to Social Interaction: Experimental Evidence from Microfinance. *The Review of Economic Studies 80*(4), 1459–1483.

Fisman, R., D. Paravisini, and V. Vig (2017). Cultural Proximity and Loan Outcomes. *The American Economic Review 107*(2), 457–492.

Franklin, S. (2016). Location, Search Costs and Youth Unemployment: Experimental Evidence from Transport Subsidies. *The Economic Journal*.

Gächter, S. and A. Falk (2002). Reputation and Reciprocity: Consequences for the Labour Relation. *Scandinavian Journal of Economics 104*(1), 1–26.

Galenianos, M. (2020). Referral Networks and Inequality. *The Economic Journal*.

Galeotti, A. and L. P. Merlino (2014). Endogenous Job Contact Networks. *International Economic Review 55*(4), 1201–1226.

Granovetter, M. S. (1973). The Strength of Weak Ties1. *American Journal of Sociology 78*(6), 1360–1380.

Heath, R. (2018). Why Do Firms Hire Using Referrals? Evidence from Bangladeshi Garment Factories. *Journal of Political Economy 126*(4), 1691–1746.

Hellerstein, J., M. Kutzbach, and D. Neumark (2014). Do Labor Market Networks Have an Important Spatial Dimension? *Journal of Urban Economics 79*, 39–58.

Hellerstein, J. K., M. McInerney, and D. Neumark (2011). Neighbors and Coworkers: The Importance of Residential Labor Market Networks. *Journal of Labor Economics 29*(4), 659–695.

Hipp, J. R., R. W. Faris, and A. Boessen (2012). Measuring 'neighborhood': Constructing Network Neighborhoods. *Social Networks 34*(1), 128–140.

Hjort, J. (2014). Ethnic Divisions and Production in Firms. *The Quarterly Journal of Economics 129*(4), 1899–1946.

Ioannides, Y. and L. Datcher Loury (2004). Job Information Networks, Neighborhood Effects, and Inequality. *Journal of Economic Literature 42*(4), 1056–1093.

Kramarz, F. and O. N. Skans (2014). When Strong Ties Are Strong: Networks and Youth Labour Market Entry. *The Review of Economic Studies 81*(3), 1164–1200.

Kugler, A. D. (2003). Employee Referrals and Efficiency Wages. *Labour Economics 10*(5), 531–556.

Magruder, J. R. (2010). Intergenerational Networks, Unemployment, and Persistent Inequality in South Africa. *AEJ: Applied Economics 2*(1), 62–85.

Montgomery, J. D. (1991). Social Networks and Labor-Market Outcomes: Toward an Economic Analysis. *The American Economic Review 81*(5), 1408–1418.

Munshi, K. (2003). Networks in the Modern Economy: Mexican Migrants in the U. S. Labor Market. *The Quarterly Journal of Economics 118*(2), 549–599.

Munshi, K. and M. Rosenzweig (2006). Traditional Institutions Meet the Modern World: Caste, Gender, and Schooling Choice in a Globalizing Economy. *American Economic Review 96*(4), 1225–1252.

Putnam, R. D. (2000). Bowling Alone: America's Declining Social Capital. In *Culture and politics*, pp. 223–234. Springer.

Rigol, N., R. Hussam, and B. Roth (2020). Targeting High Ability Entrepreneurs using Community Information: Mechanism Design in the Field. Working paper.

Sapin, M., D. Joye, C. Wolf, J. Andersen, Y. Bian, A. Carkoglu, Y.-C. Fu, E. Kalaycioglu, P. V. Marsden, and T. W. Smith (2020). The issp 2017 social networks and social resources module. *International Journal of Sociology 50*(1), 1–25.

Schmutte, I. M. (2015). Job Referral Networks and the Determination of Earnings in Local Labor Markets. *Journal of Labor Economics 33*(1), 1–32.

Serneels, P. (2007). The Nature of Unemployment among Young Men in Urban Ethiopia. *Review of Development Economics 11*(1), 170–186.

Topa, G. (2001). Social Interactions, Local Spillovers and Unemployment. *The Review of Economic Studies 68*(2), 261–295.

Topa, G. (2011). Labor Markets and Referrals. Volume 1 of *Handbook of Social Economics*, pp. 1193–1221. Elsevier North Holland Amsterdam.

Vasilaky, K. N. and K. Leonard (2018). As Good as the Networks They Keep? Improving Outcomes through Weak Ties in Rural Uganda. *EDCC*.

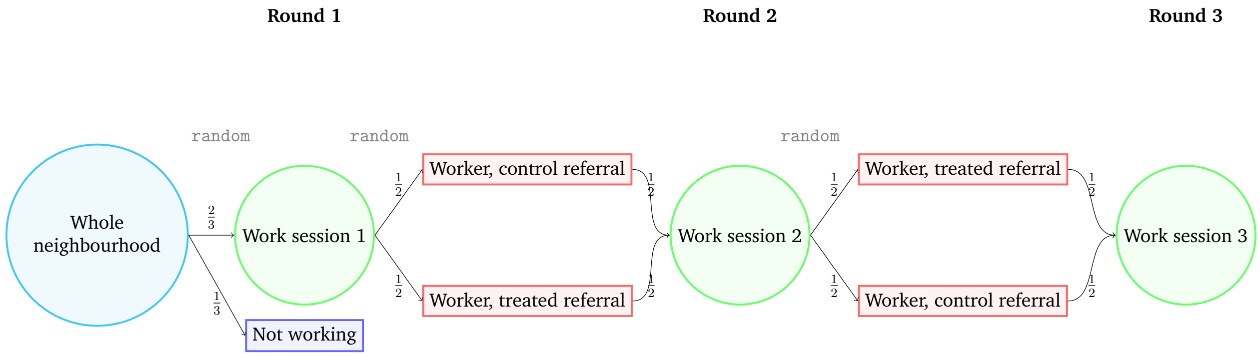
Wahba, J. and Y. Zenou (2005). Density, Social Networks and Job Search Methods: Theory and Application to Egypt. *J. of Development Economics 78*(2), 443–473.

The World Bank (2016). *World Development Indicators 2016*. Washington, DC.

Online Appendix

A Experimental design

Figure A1: Main experiment structure



B Extensions of the model

B.1 Continuous ability distribution

This extension of the model assumes a continuous distribution of the ability parameter θ_j. Each worker and referral can perform on the job with continuous ability $\theta_i \in [0, 1]$. As before, every worker i can directly observe her contact's ability θ_j.

Similar to the binary ability version of the model, every worker i expects monetary pay-offs from the employer for referring contact j as a function of the contract type P_i and the referral's ability θ_j. Workers also expect a social payoff σ_{ij} from the contact j. When selecting a contact under an incentivized referral scheme, i will receive the payoff $\Pi_{ij} = P_i(\theta_j) + \sigma_{ij}$. The referral payoff with P_i being linear in j's output is thus: $\Pi_{ij} = p_i\theta_j + \sigma_{ij}$.

Assuming that every worker i among her c_i contacts has variation in the contact j's ability, the referral decision of the worker is $\arg\max_j(\Pi_{ij} = p_i\theta_j + \sigma_{ij})$. Under referral performance incentives from the employer (P_i), workers i can maximise their payoff by selecting the referral with the highest ability θ_j:

$$E[\theta_{j^*} | j^* \in \arg\max_j(p_i\theta_j + \sigma_{ij})] \tag{A1}$$

where $\Pi_i = p_i\theta_j + \sigma_{ij}$ is the referral payoff that i receives from referring another person j for a job, with σ_{ij} being the social payoff and p_i being a monetary payoff that is linear in j's ability θ_j. The expectation is the average over all i's possible choices of j. θ_{j^*} can only change if p_i changes the choice of j^*.

In the following, I show that expression A1 is increasing in p_i.

Proof:

θ and σ are drawn from some continuous joint distribution. Under performance incentives $p_{i0} > 0$, i chooses a j_0 that maximises the payoff Π_i.

Now we introduce a higher performance payment, so that $p_{i1} > p_{i0}$. If one fixes a particular worker i and thinks about the choice that i makes under the new payment p_{i1}, on average there will be instances where i chooses the same person j_0, and instances where i will choose another person j_1.

In case of switchers under p_{i1}, the payoff for j_1 must be larger than the payoff for j_0.

$$p_{i1}\theta_{j_1} + \sigma_{ij_1} > p_{i1}\theta_{j_0} + \sigma_{ij_0} \tag{A2}$$

And under the old p_{i0}, the payoff for j_0 must be larger than the payoff for j_1.

$$p_{i0}\theta_{j_1} + \sigma_{ij_1} < p_{i0}\theta_{j_0} + \sigma_{ij_0} \tag{A3}$$

Combining A2 and A3 and re-arranging yields:

$$p_{i1}(\theta_{j_1} - \theta_{j_0}) > \sigma_{ij_0} - \sigma_{ij_1} > p_{i0}(\theta_{j_1} - \theta_{j_0}) \tag{A4}$$

As a consequence:

$$\begin{gathered} p_{i1}(\theta_{j_1} - \theta_{j_0}) > p_{i0}(\theta_{j_1} - \theta_{j_0}) \\ p_{i1}(\Delta\theta) > p_{i0}(\Delta\theta) \end{gathered} \tag{A5}$$

Since $p_{i1} > p_{i0}$, we end up with:

$$\Delta\theta > 0$$

So if i makes a different choice under p_{i1} compared to p_{i0}, it must be the case that the θ_{j1} of the j_1 selected under p_{i1} is higher than θ_{j0} of the j_0 selected under p_{i0}.

As a consequence, proposition 1 changes to:

Prop. A1: An increase in P_i increases the benefits of referring a contact with higher ability (equation A1).

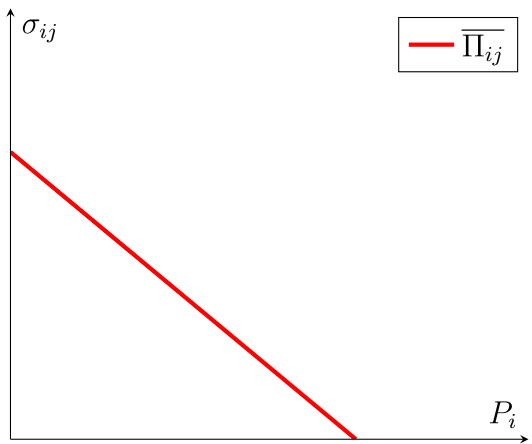
What I expect to observe in the experiment is a higher average ability of workers invited under performance incentives, as expressed by better job performance.

With the help of the opposite reasoning, it can be shown that the expected social payoff

$$E[\sigma_{ij^*} | j^* \in \arg\max_j (p_i\theta_j + \sigma_{ij})] \tag{A6}$$

is decreasing in p_i.

The referral payoff $\Pi_{ij} = P_i\theta_j + \sigma_{ij}$ demonstrates that the social payoff σ_{ij} and the performance incentive P_i are substitutes. In particular, referral contracts that include a performance incentive should *ceteris paribus* make workers i select referrals j with lower social payoffs σ_{ij}. Graphically, this could be expressed as shown below for a given referral payoff Π_{ij}.



B.2 Workers observe signals of ability

In this version of the model, I assume that the worker i cannot observe her contact's ability directly, but only receives a signal of her contact's ability $\hat{\theta}_i \in \{\theta_i^H, \theta_i^L\}$. The signal is accurate with probability β_i, where $P(\theta = \theta^H | \hat{\theta} = \theta^H, j) = P(\theta = \theta^L | \hat{\theta} = \theta^L, j) = \beta_i$, so that $\beta_i \in [0.5, 1]$. As the subscript i indicates, β_i can differ among workers.[49]

Now, worker i's monetary payoffs for referring contact j are a function of the contract type P_i, the referral's ability θ_j and the accuracy β_i of the signal $\hat{\theta}$. As before, the worker also expects a social payoff σ_{ij}. The decision between a low-ability and high-ability referral is $\sigma_{i1} \in \arg\max_j(\sigma_{ij} | \hat{\theta}_j = \theta_j^H)$ for the high-ability type and $\sigma_{i2} \in \arg\max_j(\sigma_{ij} | \hat{\theta}_j = \theta_j^L)$ for the low-ability type.

When selecting the high-signal contact under an incentivized referral scheme, i will receive the payoff $\beta_i P_i + \sigma_{i1}$, whereas for selecting the low-signal type she will

[49]The signal could actually vary for every worker-referral pair: β_{ij}.

receive $(1-\beta_i)P_i + \sigma_{i2}$. Assuming that every worker i among her contacts c_i has a choice between the high-signal and the low-signal type, the worker compares the best high-signal type with the best low-signal type. As a consequence, worker i will make a referral to a high-signal type if:

$$P_i > \frac{\sigma_{i2} - \sigma_{i1}}{2\beta_i - 1}. \tag{A7}$$

This means that the performance incentive P_i has to exceed the difference in social benefits, adjusted for the probability that the signal is accurate. The remainder of the model is as presented in section 2.

C Network summary statistics

C.1 Network level

Table C1: neighborhood network summary statistics

	(1)				
General networks	mean	min	max	sd	count
Individuals	46.19	27.00	61.00	8.28	16
Arcs	355.06	33.00	884.00	247.87	16
Density	0.16	0.05	0.43	0.10	16
In-degree centralisation	0.28	0.11	0.50	0.12	16
Out-degree centralisation	0.31	0.12	0.56	0.11	16
Transitivity	0.61	0.36	0.99	0.16	16
Reciprocity	0.27	0.17	0.43	0.08	16
Average between centrality	42.24	1.30	73.41	17.98	16
Average degree centrality	9.31	1.33	19.24	4.73	16
Average eigenvector centrality	0.12	0.11	0.14	0.01	6
Average Katz centrality	21.60	14.88	35.11	5.91	16
Average Clustering coefficient	0.37	0.14	0.58	0.13	16
Number of components	2.88	1.00	10.00	2.66	16
Job networks	mean	min	max	sd	count
Individuals	46.19	27.00	61.00	8.28	16
Arcs	102.88	11.00	301.00	82.09	16
Density	0.05	0.02	0.15	0.03	16
In-degree centralisation	0.13	0.06	0.21	0.05	16
Out-degree centralisation	0.17	0.03	0.49	0.12	16
Transitivity	0.75	0.21	1.27	0.31	15
Reciprocity	0.31	0.10	0.52	0.10	16
Average between centrality	13.45	0.04	47.26	15.34	16
Average degree centrality	2.13	0.41	6.57	1.62	16
Average Katz centrality	39.64	26.77	51.33	6.75	16
Average Clustering coefficient	0.19	0.00	0.40	0.10	16
Number of components	14.44	3.00	24.00	6.92	16

Notes: The variables labelled are the means of individuals within a network, of which subsequently is taken an average across networks. Centralisation scores are divided by N - 1, where N = number of nodes in a network. This standardization makes sure that centrality scores always range from 0 to 1.

C.2 By network

Table C2: General and job network-level summary and centrality statistics

General networks:

#	betweenness	degree	eigenvector	katz	clustering	# of components	nodes	arcs	density	in-degree centralisation	out-degree centralisation	transitivity	reciprocity
1	25.72	19.24	0.140	18.27	0.576	1	46	884	0.427	0.358	0.562	0.995	0.433
2	21.96	3.362	.	35.11	0.215	10	47	158	0.0731	0.170	0.236	0.619	0.206
3	37.51	3.404	.	30.74	0.305	2	47	159	0.0735	0.235	0.124	0.463	0.339
4	50.82	4.245	.	26.53	0.282	5	49	208	0.0884	0.144	0.293	0.608	0.216
5	44.11	2.711	.	20.53	0.140	2	38	103	0.0733	0.147	0.202	0.394	0.212
6	34.44	6.917	.	14.94	0.508	2	36	249	0.198	0.267	0.267	0.362	0.415
7	29.88	3.605	.	27.71	0.225	3	43	154	0.0853	0.205	0.278	0.560	0.225
8	51.12	11.84	0.125	19.16	0.474	1	56	662	0.215	0.336	0.410	0.601	0.304
9	57.94	8.815	.	19.19	0.418	2	54	475	0.166	0.503	0.273	0.544	0.219
10	64.49	7.667	0.120	15.45	0.440	1	51	391	0.153	0.456	0.313	0.687	0.192
11	52.82	4.644	.	17.66	0.348	3	45	209	0.106	0.264	0.334	0.502	0.215
12	73.41	9.902	0.108	19.67	0.436	1	61	604	0.165	0.290	0.493	0.657	0.277
13	1.296	1.333	.	25.63	0.175	8	27	33	0.0470	0.107	0.186	0.844	0.309
14	55.76	5.643	0.131	14.88	0.337	1	42	237	0.138	0.209	0.434	0.564	0.167
15	41.79	12.70	0.120	21.59	0.495	1	53	672	0.244	0.339	0.280	0.527	0.369
16	32.84	10.98	.	18.58	0.483	3	44	483	0.255	0.429	0.310	0.786	0.291

Job networks:

#	betweenness	degree	eigenvector	katz	clustering	# of components	nodes	arcs	density	in-degree centralisation	out-degree centralisation	transitivity	reciprocity
1	27.37	6.565	.	28.82	0.231	3	46	301	0.145	0.214	0.487	1.063	0.277
2	1.851	1.170	.	45.31	0.124	18	47	55	0.0254	0.0851	0.129	0.260	0.279
3	3.468	1.447	.	44.54	0.241	22	47	68	0.0315	0.101	0.168	0.964	0.388
4	0.388	1.102	.	48.18	0.156	24	49	54	0.0230	0.0829	0.0829	0.844	0.421
5	2.289	0.789	.	36.74	0.105	18	38	30	0.0213	0.0614	0.0336	0.214	0.364
6	10.97	2.611	.	30.88	0.404	5	36	94	0.0746	0.158	0.129	0.545	0.516
7	1.256	1.023	.	41.81	0.116	23	43	43	0.0238	0.146	0.121	0.945	0.397
8	41.50	2.589	.	39.70	0.247	7	56	145	0.0471	0.0817	0.211	0.914	0.318
9	47.26	2.907	.	36.90	0.265	6	54	156	0.0545	0.175	0.213	0.524	0.231
10	16.71	1.961	.	42.91	0.0887	13	51	100	0.0392	0.184	0.246	0.443	0.163
11	1	0.756	.	44.04	0.0815	22	45	34	0.0172	0.0754	0.0522	1.125	0.259
12	23.23	2.574	.	51.33	0.286	16	61	157	0.0429	0.160	0.194	1.271	0.319
13	0.0370	0.407	.	26.77	0	18	27	11	0.0157	0.0636	0.0636	.	0.100
14	5.476	1.595	.	38.66	0.205	16	42	67	0.0389	0.0851	0.160	0.857	0.340
15	26.26	4.887	.	37.23	0.297	8	53	258	0.0936	0.198	0.355	0.640	0.332
16	6.091	1.659	.	40.41	0.157	12	44	73	0.0386	0.151	0.127	0.694	0.259

C.3 Within network

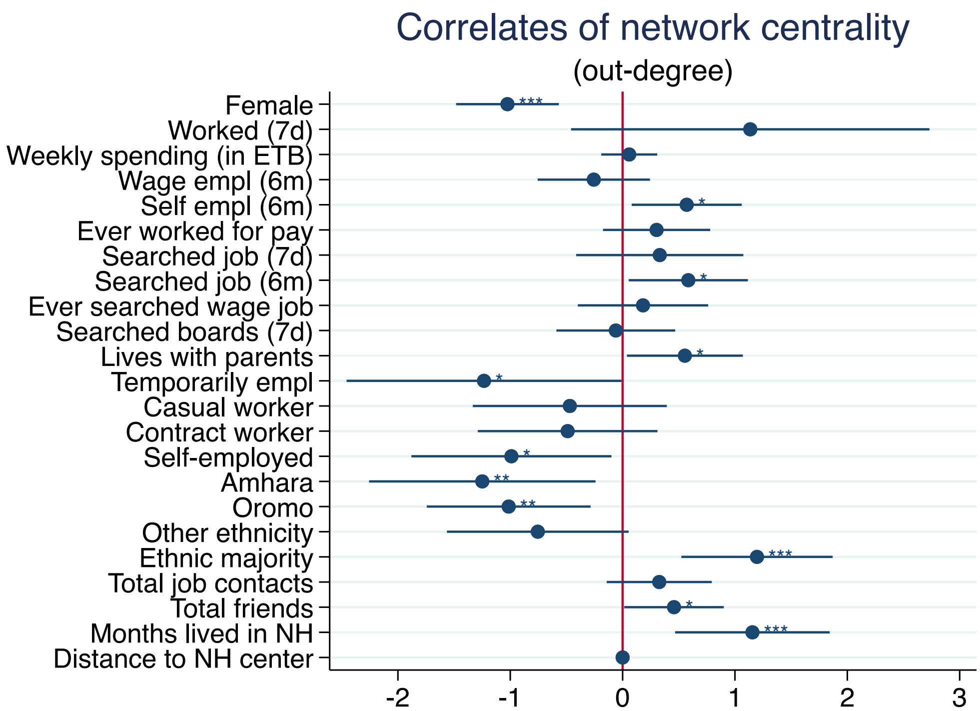
Table C3: Connections within the neighborhood networks (shares)

Share of total network:	mean	min	max	sd	count
All links	0.17	0.00	0.96	0.15	739
Considered information spreaders	0.12	0.00	8.32	0.38	739
Travel connections	0.04	0.00	0.43	0.06	739
Links visited	0.07	0.00	0.55	0.09	739
Visitors	0.07	0.00	0.57	0.09	739
Job info given to	0.05	0.00	0.53	0.07	739
Job info received from	0.04	0.00	0.58	0.07	739
Money lent to	0.03	0.00	0.34	0.05	739
Money borrowed from	0.03	0.00	0.34	0.04	739

Notes: Table C3 displays the summary statistics of the social networks within each of the 16 neighbourhoods, dividing the number of connections of a certain type of individual i by the overall network size.

D Summary statistics, take-up and covariate balance

Figure A2: Regression of out-degree centrality on individual covariates



Notes: Figure A2 displays the standardised regression coefficients from a multivariate OLS regression of individual out-degree centrality on individual characteristics. NH=neighborhood. $N = 739$. Imputed missing values for weekly spending (ETB) and number of months lived in NH. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level.

Table D4: Balance of covariates for baseline sample, by round one treatment status

	(1) Referral	(2) Mean of control (SD)	(3) Max pairwise difference	(4) Obs.
Age	0.05	23.78	0.01	507
	(0.31)	(3.43)		
Female	0.06	0.43	0.12	513
	(0.04)	(0.50)		
Worked (7d)	0.02	0.20	0.04	513
	(0.04)	(0.40)		
# of friends	2.14	8.52	0.25	513
	(1.31)	(7.64)		
Wage empl (6m)	0.10*	0.31	0.20	513
	(0.05)	(0.46)		
Self empl (6m)	-0.03	0.20	0.07	513
	(0.03)	(0.40)		
Ever worked for pay	0.03	0.78	0.07	513
	(0.03)	(0.42)		
Searched job (7d)	-0.01	0.40	0.01	513
	(0.04)	(0.49)		
Searched job (6m)	0.02	0.66	0.04	513
	(0.05)	(0.47)		
Ever searched wage job	0.03	0.49	0.05	513
	(0.03)	(0.50)		
Searched boards (7d)	0.03	0.19	0.07	513
	(0.03)	(0.39)		
Lives with parents	0.06	0.60	0.13	513
	(0.06)	(0.49)		
Years since school	7.67	2.10	0.10	503
	(7.20)	(3.62)		
Permanently empl	0.00	0.00	0.00	513
	(0.01)	(0.06)		
Temporarily empl	0.01	0.08	0.03	513
	(0.02)	(0.27)		
Casual worker	-0.00	0.03	0.01	513
	(0.02)	(0.18)		
Contract worker	0.03	0.03	0.13	513
	(0.02)	(0.16)		
Self-employed	-0.01	0.04	0.03	513
	(0.02)	(0.20)		
Weekly spending (in ETB)	73.62	432.94	0.11	491
	(52.50)	(538.63)		
Amhara	-0.01	0.39	0.01	513
	(0.07)	(0.49)		
Oromo	0.01	0.22	0.03	513
	(0.04)	(0.42)		
Other ethnicity	-0.01	0.32	0.03	513
	(0.04)	(0.47)		

Notes: OLS estimates of individual baseline differences of treatment groups. Outcome variables are listed on the left. Standard errors are in parentheses. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level. All monetary values are displayed as converted from Ethiopian birr (ETB) to 2015 USD, with an exchange rate of 21.5 ETB per 1 USD, the average rate from December 2015 to March 2016. In column 3, I calculate the pairwise difference between the two group means and divide this by the standard deviation of the variable, following Imbens (2015).

Table D5: Balance of covariates for baseline sample, by attendance

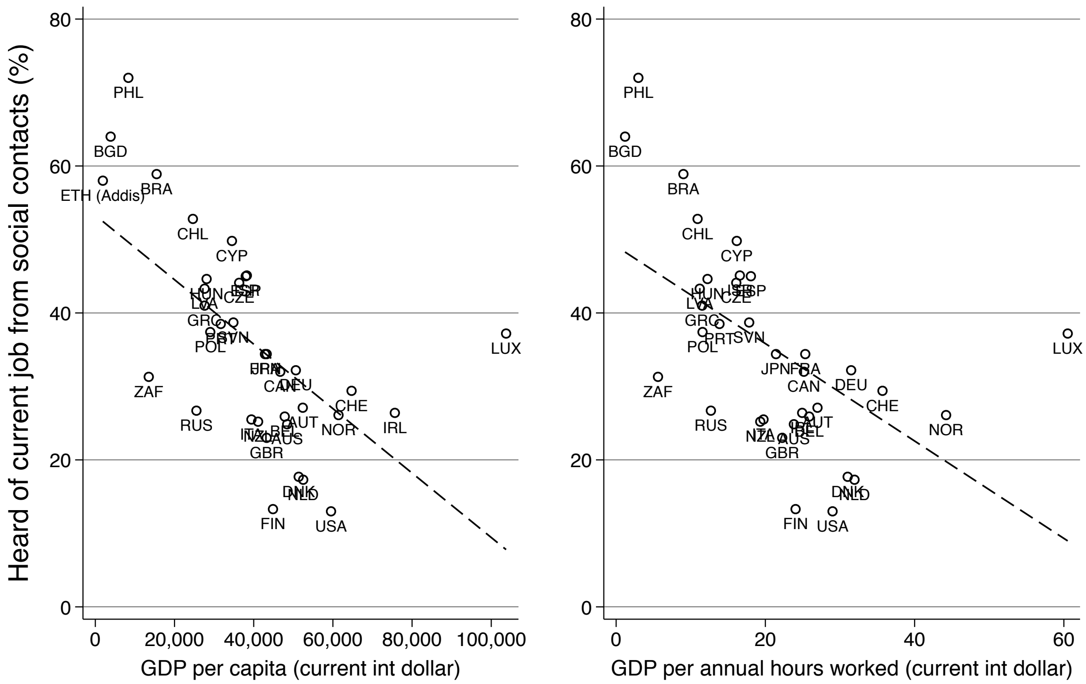
	(1) Attended	(2) Mean of non-attenders (SD)	(3) Max pairwise difference	(4) Obs.
Age	-0.38	24.18	0.11	729
	(0.32)	(3.45)		
Female	-0.03	0.49	0.06	739
	(0.06)	(0.50)		
Worked (7d)	-0.08*	0.29	0.20	739
	(0.05)	(0.45)		
# of friends	2.68***	6.88	0.32	739
	(0.73)	(7.19)		
Wage empl (6m)	-0.06	0.42	0.13	739
	(0.05)	(0.49)		
Self empl (6m)	0.02	0.17	0.06	739
	(0.03)	(0.37)		
Ever worked for pay	-0.02	0.81	0.05	739
	(0.03)	(0.39)		
Searched job (7d)	0.09***	0.30	0.19	739
	(0.03)	(0.46)		
Searched job (6m)	0.08*	0.58	0.18	739
	(0.05)	(0.49)		
Ever searched wage job	-0.12***	0.63	0.24	739
	(0.04)	(0.48)		
Searched boards (7d)	0.06**	0.15	0.15	739
	(0.02)	(0.35)		
Lives with parents	0.11**	0.52	0.23	739
	(0.05)	(0.50)		
Years since school	3.41	2.39	0.05	724
	(3.63)	(3.89)		
Permanently empl	-0.00	0.00	0.01	739
	(0.01)	(0.07)		
Temporarily empl	-0.07*	0.15	0.22	739
	(0.03)	(0.36)		
Casual worker	0.00	0.03	0.01	739
	(0.02)	(0.17)		
Contract worker	-0.01	0.05	0.05	739
	(0.02)	(0.22)		
Self-employed	-0.01	0.04	0.03	739
	(0.01)	(0.21)		
Weekly spending (in ETB)	-181.80	650.88	0.14	702
	(169.33)	(2170.03)		
Amhara	-0.04	0.43	0.08	739
	(0.04)	(0.50)		
Oromo	0.03	0.20	0.07	739
	(0.02)	(0.40)		
Other ethnicity	0.03	0.28	0.07	739
	(0.04)	(0.45)		

Notes: OLS estimates of individual baseline differences by attendance. Outcome variables are listed on the left. Standard errors are in parentheses. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level. All monetary values are displayed as converted from Ethiopian birr (ETB) to 2015 USD, with an exchange rate of 21.5 ETB per 1 USD, the average rate from December 2015 to March 2016. In column 3, I calculate the pairwise difference between the two group means and divide this by the standard deviation of the variable, following Imbens (2015).

E Additional results

E.1 Cross-country evidence on obtaining jobs through social networks

Figure A3: Correlation between finding jobs through social networks and GDP per capita or labor productivity



Source: data compiled from ECHP, ISSP, OurWorldInData and own data. Labour productivity is defined as GDP per capita over working hours per year, in 2013

E.2 Correlation between productivity and centrality

I test whether the tendency of workers to invite more central individuals to the day job has implications for productivity. In particular, is it the case that more central individuals are also more able, hence perform better on the job? Table E6 regresses the workers' work performance on their network centrality, for the three different rounds. It is evident that there is no statistically significant relationship between a worker's network centrality and her work performance. In order to disentangle the effects of referring more productive workers who also *happen* to be more central from work productivity itself, columns (3) and (5) look at workers invited by the lottery/computer: the correlation between centrality and productivity actu-

Table E6: Pooled cross-sections: Central individuals are not more productive

	Round 1	Round 2		Round 3	
	(1)	(2)	(3)	(4)	(5)
	Work performance$_j$	Work performance$_j$	Work performance$_{randomj}$	Work performance$_j$	Work performance$_{randomj}$
Indegree centrality$_j$	0.0109	0.0450	-0.0106	0.00808	-0.117
	(0.26)	(1.02)	(-0.09)	(0.21)	(-1.38)
Constant	10.66***	10.71***	9.986***	11.51***	11.96***
	(20.58)	(20.37)	(9.40)	(24.82)	(14.82)
N	261	270	80	270	78

Notes: The work scores display the first-session performance of the workers, which directly captures individual work ability rather than learning on the job or other repeated work considerations. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level.

ally becomes negative here, but remains statistically insignificant. I interpret this as strong evidence that central workers are not more productive — so the fact that more central individuals are more likely to get a job has no positive implications for productivity.

E.3 Robustness to shrouding

E.3.1 Additional treatment variation: centrality shrouding

Since I investigate the implications of worker referrals on labor market equality, one potential test is whether participants refer workers in their social network who have a higher degree centrality, i.e. more links to other nodes. In reality, an individual's degree centrality is not randomly assigned and likely correlates with other hard-to-observe or unobservable personal characteristics (e.g. sociability, physical attractiveness, intelligence). In order to distinguish between such personal traits and the isolated effects of a person's connectedness, I introduce a separate exogenous sub-treatment based on between-worker variation. In some neighborhoods, a random half of participants invites a worker based on a referral list where the true number of connections (i.e. the degree centrality) is shrouded. The remaining workers get the usual referral lists, i.e. stating the number of connections explicitly for each person j in the neighborhood. In practice, the referral with the shrouded degree centrality means that the referral list contains two numbers of friends for each potential referral j, one true number and one random false number. As I elicit every person i's belief about j's number of friends at baseline, I control for these baseline beliefs in my regressions in section 5.

In order to test whether underlying characteristics of central workers determine their job prospects rather than their pure network centrality, I introduce the sub-treatment that exploits the exogenous experimental variation in whether individual j's degree centrality was made salient to the workers or not (i.e. the centrality shrouding sub-treatment).[50]

Table E7 shows the results. I find that those workers j who were invited through job referrals that made their network centrality salient (revealed degree centrality on the referral lists) have on average three more connections than the average network member. Workers j who were invited with concealed job referrals (i.e. their degree centrality was not a salient characteristic during the referral process) are also more central than the average network member, with on average two more connections (which is statistically not significant). The comparisons to the average network member are based on non-random variation, as both types of referrals induce selection. On the other hand, the comparison of the coefficients on C and T4 is identified through experimental variation. The difference of one connection on average between the two treatments may bear some economic significance, but I cannot reject that the coefficients for the two different treatments (revealed vs shrouded degree centrality) are equal ($p = 0.695$); in fact, they seem reasonably close to each other. This supports an interpretation where network centrality bears no additional meaning beyond the individual characteristics that are associated with it. The additional value individuals receive from referring more central individuals to the work sessions can be almost fully explained by their innate characteristics, rather than by their pure position in the network alone.

[50] I pool the experimental data to run the following regression:

$$Centrality_j = \alpha_0 + \alpha_1 \cdot Referral_i^{Open} + \alpha_2 \cdot Referral_i^{Shrouded} + \theta X_{ij} + u_j, \quad \text{(A8)}$$

where $Referral_j^{Shrouded}$ indicates a dummy for the shrouded degree centrality treatment. I test the hypothesis $\alpha_1 = \alpha_2$ to detect if degree centrality has a separate effect beyond correlated confounders.

Table E7: Pooled cross-sections: Open vs shrouded degree centrality referral

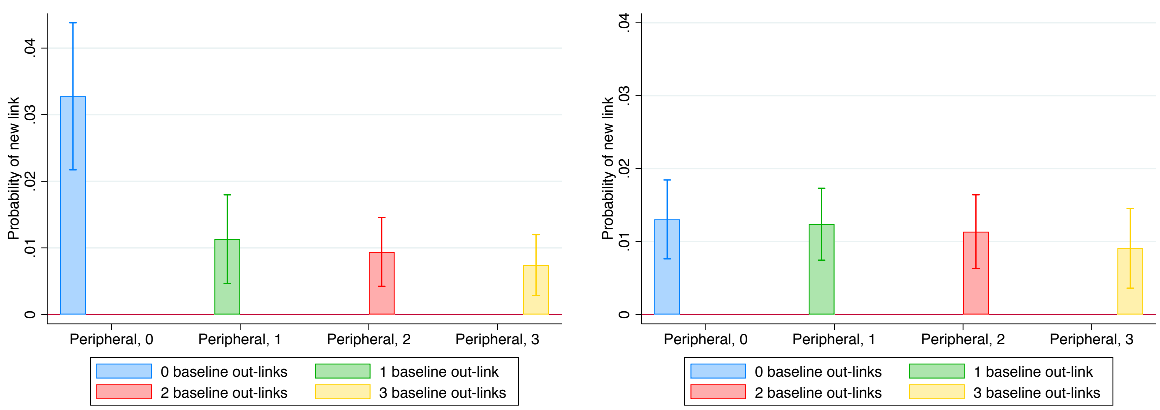
	Degr. Centr.$_j$
i invited j with standard referral (C)	2.923**
	(1.17)
i invited j with concealed referral (T4)	2.019
	(1.41)
Constant	9.838***
	(0.63)
Number of observations	480
p-value for H0: Degr. Centr.$_j$ (open referral)=Degr. Centr.$_n$ (concealed referral)	0.695

Notes: OLS regression of j's degree centrality on the (exogenous) type of referral treatment. Standard errors in parentheses. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level.

E.4 Periphery thresholds

Figure A4: Dyadic regressions: Different periphery thresholds

(a) General networks (b) Job networks

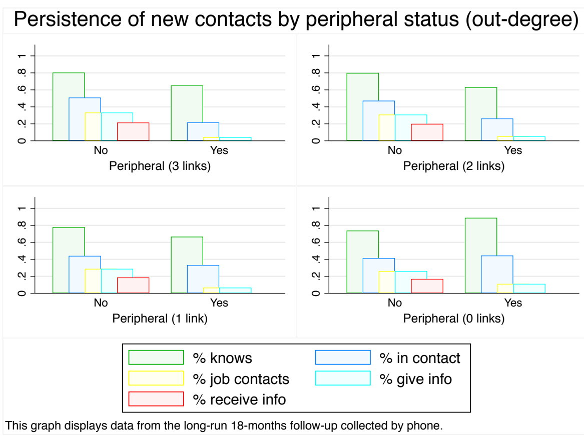


E.5 Long-run effects

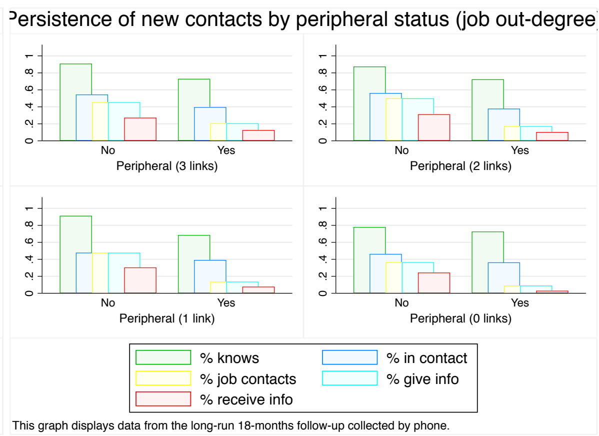
Figure A5 displays the number of newly established links that were still in some form of relation 18 months after the intervention, by whether the worker was peripheral at baseline for different periphery threshold. The top figures defines periphery based on the worker's in-degree centrality, the bottom ones based on out-degree. The left figures look at general networks, the right at job networks. Two things become apparent in the graphs. First, peripheral workers have lower social interactions across all categories than non-peripheral workers, regardless of the threshold chosen to define peripheral individuals. Second, while this finding holds across all panels it is stronger for in-degree than for out-degree, and for general networks than for job networks. In particular, peripheral workers in job networks often manage to receive job information to the same extent as non-peripheral workers. They also receive information to the same extent than they give, which is not true for non-peripheral workers.

Figure A5: Long-run persistence of new links by baseline peripheral status

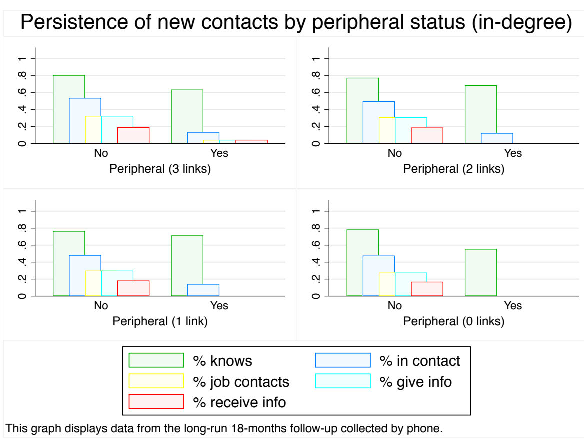
(a) General networks – out-degree



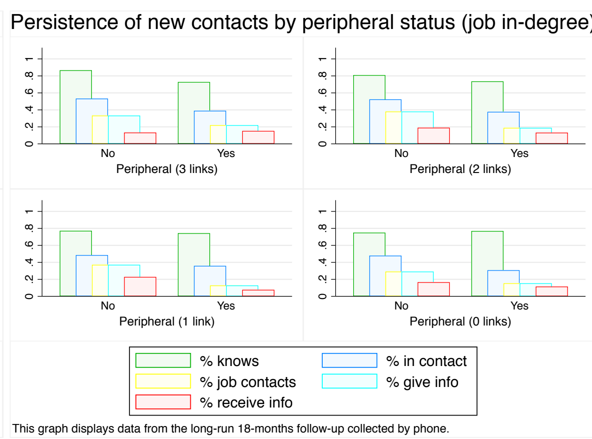
(b) Job networks – out-degree



(c) General networks – in-degree



(d) Job networks – in-degree



E.6 Correlation between centrality measures and getting a referral

The second panel of table E8 shows results for a regression of a binary indicator of whether individual i referred individual j to the job on a variety of job network

Table E8: Pooled cross-sections in general and job networks: More central individuals are referred to the job

General networks	Referral_{ij}	Referral_{ij}	Referral_{ij}	Referral_{ij}	Referral_{ij}	Referral_{ij}	Referral_{ij}	Referral_{ij}	Referral_{ij}
Outdegree centrality_j	0.0100^{***}		0.00410						0.00593^{*}
	(4.62)		(1.42)						(1.96)
Indegree centrality_j		0.0120^{***}	0.00920^{***}						0.0101^{***}
		(5.40)	(3.09)						(3.15)
Between centrality_j				0.0000304					-0.000588^{**}
				(0.13)					(-2.16)
Out-Katz centrality_j					-0.000691				0.00119
					(-0.40)				(0.66)
In-Katz centrality_j						-0.00152			-0.000473
						(-0.93)			(-0.28)
Clustering centrality_j							0.140^{**}		-0.00628
							(2.38)		(-0.09)
Eigenvector centrality_j								0.454	
								(1.14)	
Constant	0.346^{***}	0.327^{***}	0.316^{***}	0.431^{***}	0.448^{***}	0.464^{***}	0.375^{***}	0.482^{***}	0.309^{***}
	(13.63)	(12.60)	(11.69)	(20.38)	(10.89)	(12.34)	(12.67)	(8.62)	(4.73)
N	836	836	836	836	836	836	836	406	836
Job networks									
Outdegree centrality_j	0.0123^{***}		0.00295					0.00382	
	(3.13)		(0.57)					(0.66)	
Indegree centrality_j		0.0208^{***}	0.0183^{***}					0.0151^{**}	
		(4.13)	(2.73)					(1.99)	
Between centrality_j				0.000585^{*}				-0.000304	
				(1.88)				(-0.83)	
Out-Katz centrality_j					-0.00701^{***}			-0.00485^{**}	
					(-4.16)			(-2.52)	
In-Katz centrality_j						-0.00523^{***}		-0.00290	
						(-3.63)		(-1.56)	
Clustering centrality_j							-0.0218	-0.163^{***}	
							(-0.44)	(-2.94)	
Constant	0.397^{***}	0.373^{***}	0.372^{***}	0.420^{***}	0.712^{***}	0.640^{***}	0.439^{***}	0.734^{***}	
	(19.25)	(16.74)	(16.55)	(22.85)	(10.29)	(10.74)	(20.66)	(6.73)	
N	836	836	836	836	836	836	836	836	

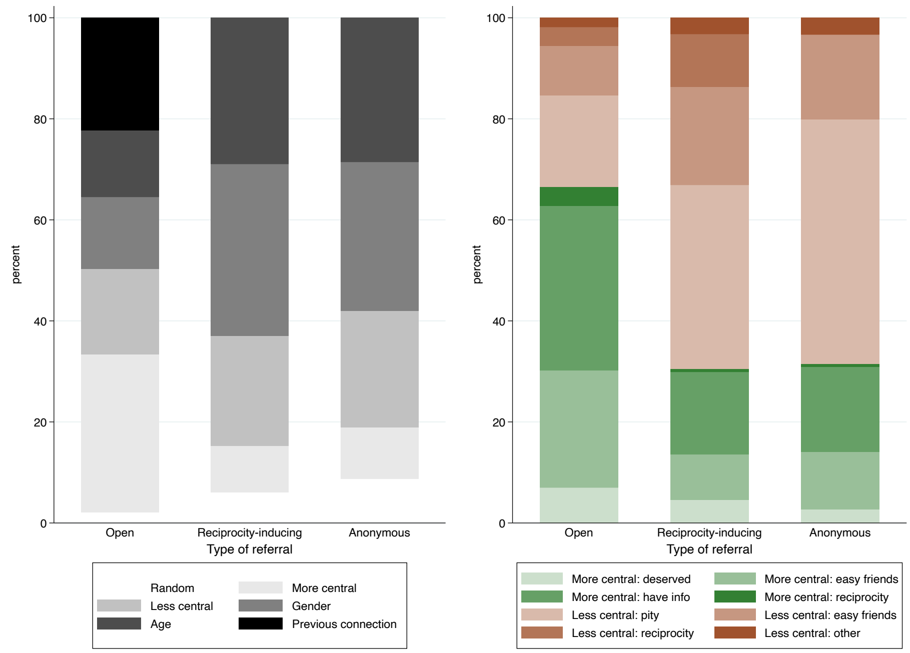
Notes: OLS regression of a binary indicator indicating whether individual i referred individual j to the job on individual j's centrality in general and job networks. The regressions are run on the pooled cross-sectional data (i.e. one observation per individual). * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level.

centrality measures. The positive effects found in general social networks for in- and out-degree centrality persist, with in-degree dominating again. If individual j is in the job information network of one more person, her chances of being referred to the job increase by almost 6%. Interestingly, when all network centrality measures enter the regression simultaneously, the individual clustering coefficient enters as a strong negative predictor of the chances of receiving a job referral.[51] An intuitive explanation of the local clustering coefficient is how close j's neighbors (*excluding* j) are to being a clique, or a complete graph, meaning that all neighbors N_i are also interconnected among themselves and not just through j. This brings more nuance to the previous finding: if j is a job information contact of individual i, j's chances of being referred to the job increase. However, if individual i is also connected to j's contacts k and l, j's chances decrease *ceteris paribus*. In other words, if j and i are a connected pair of job information partners, j has higher chances of being invited to a job than in a triplet where i, j and k are all connected (complete graph). This underlines the rivalrous aspects of job referrals: existing job information networks predict who gets the referral. In a complete clique of job information sharers, the referral can go to multiple individuals, whereas in pairs of individuals, the potential recipient is only one person – unless the referee links up with a new node.

E.7 Self-reported reasons behind a particular referral

[51]The local clustering coefficient of node j in the network is defined as the share of network contacts N_i of j who are directly connected among themselves.

Figure A6: Workers' self-reported reasons behind a particular referral



Notes: This figure presents individual-level short-run follow-up data on self-reported reasons behind referrals. The left panel shows the percentage of respondents claiming specific motive behind a referral, by referral type. The right panel shows the self-reported motive behind those respondents who referred based on network centrality, by referral type.

E.8 Reciprocity in job networks

Table E9: Dyadic regressions: peripheral individuals benefit from reciprocity

Specification	(1)	(2)		
General networks	Referral^2_{ij}	Referral^2_{ij}	Referral^2_{ij}	Referral^2_{ij}
Peripheral_j*New Referral^1_{ji}	0.235***	0.235***	0.236***	0.234***
	(0.04)	(0.04)	(0.04)	(0.04)
Peripheral_j*Any Referral^1_{ji}	-0.0559	-0.0559	-0.0607*	-0.0617*
	(0.04)	(0.04)	(0.04)	(0.04)
New referral^1_{ji}	-0.0988***	-0.0988***	-0.0807***	-0.0821***
	(0.02)	(0.02)	(0.02)	(0.02)
Any Referral^1_{ji}	0.320***	0.321***	0.307***	0.307***
	(0.01)	(0.01)	(0.01)	(0.01)
Random_i		0.00470**	0.00512**	0.0103***
		(0.00)	(0.00)	(0.00)
Constant	0.0100***	0.00928***	0.0197***	0.0285***
	(0.00)	(0.00)	(0.01)	(0.01)
Baseline controls	No	No	Yes	Yes
NH fixed effects	No	No	No	Yes
Number of observations	17280	17280	17280	17280
p-value for H0:				
Any Referral^1_{ji}+New referral^1_{ji}+ Peripheral_j*New Referral^1_{ji}=0 (=Total $\text{effect}^{per}_{new}$)		0.000	0.000	0.000
Total $\text{effect}^{per}_{new}$=$\text{Random}_i$		0.000	0.000	0.000
Total $\text{effect}^{per}_{new}$=Total $\text{effect}^{per}_{any}$		0.008	0.003	0.003
Job networks				
Peripheral_j*New Referral^1_{ji}	-0.0247	-0.0247	-0.0205	-0.0178
	(0.03)	(0.03)	(0.03)	(0.03)
Peripheral_j*Any Referral^1_{ji}	0.0469	0.0469	0.0415	0.0383
	(0.03)	(0.03)	(0.03)	(0.03)
New referral^1_{ji}	0.0149	0.0149	0.0263	0.0243
	(0.02)	(0.02)	(0.02)	(0.02)
Any Referral^1_{ji}	0.261***	0.262***	0.248***	0.249***
	(0.02)	(0.02)	(0.02)	(0.02)
Random_i		0.00468**	0.00522**	0.0103***
		(0.00)	(0.00)	(0.00)
Constant	0.00960***	0.00887***	0.0175**	0.0262***
	(0.00)	(0.00)	(0.01)	(0.01)
Baseline controls	No	No	Yes	Yes
NH fixed effects	No	No	No	Yes
Number of observations	17280	17280	17280	17280
p-value for H0:				
Any Referral^1_{ji}+New referral^1_{ji}+ Peripheral_j*New Referral^1_{ji}=0 (=Total $\text{effect}^{per}_{new}$)		0.000	0.000	0.000
Total $\text{effect}^{per}_{new}$=$\text{Random}_i$		0.000	0.000	0.000
Total $\text{effect}^{per}_{new}$=Total $\text{effect}^{per}_{any}$		0.272	0.488	0.536

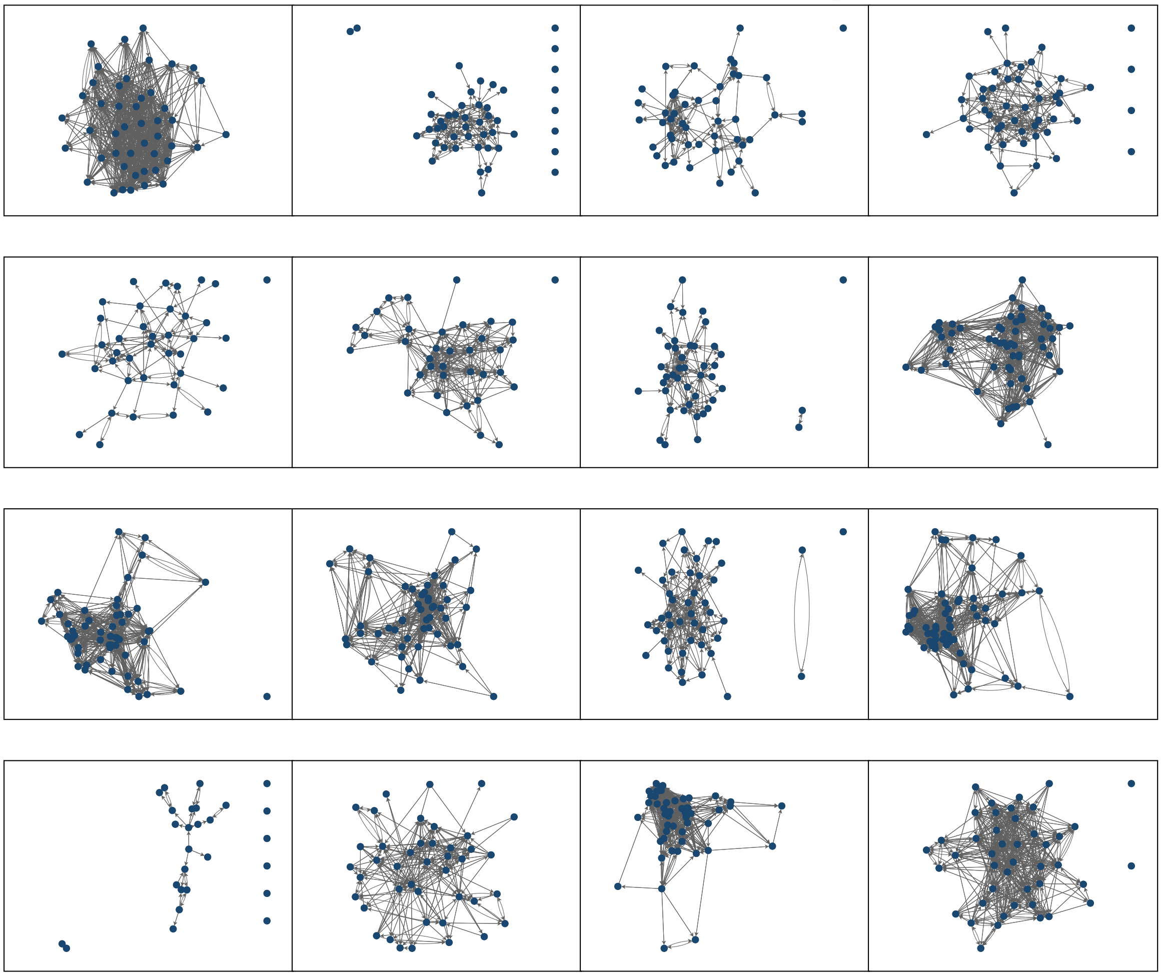
Notes: Table E9 regresses an indicator of whether worker i referred worker j after round 2 of the experiment on whether in the previous round j had invited i, plus several referral type controls and interactions with j's periphery status. Here, an individual i is defined as peripheral when having fewer than two out-degree connections. The top panel shows the results for general networks, the bottom panel for job networks. Standard errors in parentheses. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level.

Notes: Table E9 regresses an indicator of whether worker i referred worker j after round 2 of the experiment on whether in the previous round j had invited i, plus several referral type controls and interactions with j's periphery status. Here, an individual i is defined as peripheral when having fewer than two out-degree connections. Standard errors in parentheses. * denotes significance at 10 pct., ** at 5 pct., and *** at 1 pct. level.

F Network maps

F.1 General networks

Figure A7: Overview of directed social network graphs at baseline



Notes: This figure presents the directed social network graphs for all 16 neighborhoods at baseline. An arrow between two nodes represents a directed link, which is defined as an individual replying positively to the question "Do you know [$name\ of\ j$]?". The networks include all eligible resident individuals in the neighborhood, where eligible individuals 1) permanently live in the selected neighborhood, 2) are between 18 and 29 years of age (inclusive), and 3) are not in permanent employment or education.

F.2 Job networks

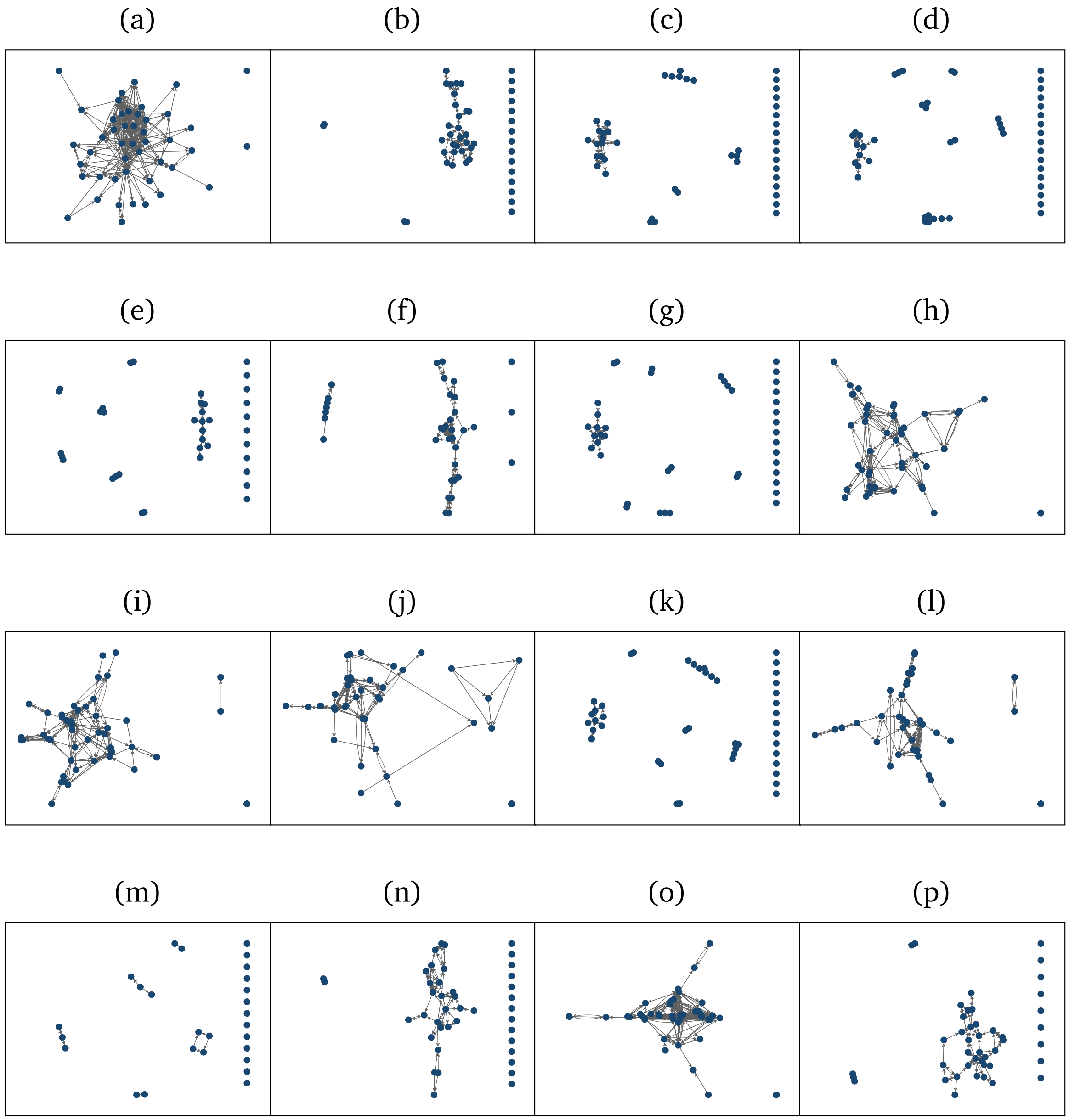


Figure A8: Overview of directed job referral network graphs at baseline