Are gifts and loans between households voluntary?

Margherita Comola, Marcel Fafchamps

PARIS SCHOOL OF ECONOMICS
ECOLE D'ECONOMIE DE PARIS

WORKING PAPER N° 2010 - 19

Are gifts and loans between households voluntary?

Margherita Comola

Marcel Fafchamps

JEL Codes: C13, C51, D85
Keywords: risk sharing, reporting bias, social networks

PARIS-JOURDAN SCIENCES ECONOMIQUES
48, BD JOURDAN – E.N.S. – 75014 PARIS
TÉL. : 33(0) 1 43 13 63 00 – FAX : 33 (0) 1 43 13 63 10
www.pse.ens.fr

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE – ECOLE DES HAUTES ETUDES EN SCIENCES SOCIALES
ÉCOLE DES PONTS PARISTECH – ECOLE NORMALE SUPÉRIEURE – INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE

Are Gifts and Loans between Households Voluntary?*

Margherita Comola† Marcel Fafchamps‡
Paris School of Economics Oxford University

June 2010

Abstract

Using village data from Tanzania, we test whether gifts and loans between households are voluntary while correcting for mis-reporting by the giving and receiving households. Two maintained assumptions underlie our analysis: answers to a question on who people would turn to for help are good proxies for willingness to link; and, conditional on regressors, the probability of reporting a gift or loan is independent between giving and receiving households. Building on these assumptions, we develop a new estimation methodology that corrects for response bias. Our testing strategy is based on the idea that, if lending and gift giving are voluntary, then both households should want to rely on each other for help. We find only weak evidence to support bilateral link formation. We do, however, find reasonably strong evidence to support unilateral link formation. Results suggest that if a household wishes to enter in a reciprocal relationship with someone who is sufficiently close socially and geographically, it can do so unilaterally.

Keywords: risk sharing; reporting bias; social networks
JEL codes: C13; C51; D85

*We are indebted to Joachim De Weerdt for sharing his data and answering our questions. We benefitted from comments from seminar participants at the Paris School of Economics and the Norwegian School of Economics and Business Administration.

†Paris School of Economics. Email: comola@pse.ens.fr
‡Department of Economics, University of Oxford. Email: marcel.fafchamps@economics.ox.ac.uk

1 Introduction

Much social network analysis is based on dyadic data reported by survey respondents (e.g., Christakis and Fowler 2009, Steglich, Snijders, and Pearson 2010, Fafchamps and Lund 2003). It is common for such data to be discordant. For instance, when asked to list their friends, it often arises that, say, i lists j as a friend but the reverse is not true. Similarly, i may report giving something to j while j does not report receiving anything from i.

Researchers typically ignore this issue even though it affects estimation and inference. To see why, consider the following example. Suppose that half of those who give report giving while a third of those who receive report receiving. Further assume that reporting errors are uncorrelated between giver and receiver. This means that answers agree in one sixth of the observations; the rest are discordant. Finally, suppose – as many researchers do – that a gift is assumed to have taken place if it is reported either by the giver or the receiver. With these assumptions, the actual number of gifts is underestimated by one third.[1]

Inference may further be biased if factors that affect giving differently affect non-response. For instance, imagine that the old and the young give to each other equally but the old are more likely to report than the young. The combined data will erroneously suggest that the old are more likely to give to the young than the opposite.

This paper proposes a methodology to deal with such mis-reporting. This methodology can be extended to other types of response bias. The theoretical literature on networks has emphasized the role that one-sided or two-sided link formation have on the equilibrium topology of social networks (Goyal 2007, Jackson 2009). Using our methodology to explicitly recognize the possibility of mis-reporting, we test whether transfers (i.e. loans and gifts) between households are more consistent with unilateral or bilateral link formation.

Two maintained assumptions underlie our analysis. The first is that answers to a first-round question on who people would turn to for help are good proxies for willingness to link. After examining these answers in detail, Comola and Fafchamps (2009) indeed conclude that they are best interpreted as self-censored willingness to link. We provide additional evidence here that it is a reasonable assumption. The second maintained hypothesis is that reporting propensities are independent between giving and receiving households. This assumption only enters into the construction of the maximum likelihood estimator that explicitly corrects for under-reporting; it does not affect simpler regression analysis that we also conduct.

Our testing strategy is based on the idea that, if lending and gift giving are a voluntary agreement between two households, then both households should want to rely on each other for help. In contrast, if households cannot refuse to assist others, then gifts and loans are best seen as a unilateral process. Results provide only weak support for bilateral link formation.

[1] Relative to the actual number of gifts, the observed number of gifts is equal to 1/2 (reported by givers) +1/3 (reported by receivers)-1/6 (reported by both) = 2/3.

We do, however, find evidence that is in agreement with the unilateral link formation hypothesis. These findings suggest that surveyed households probably find it difficult to extricate themselves from social and familial obligations to assists others in need. This stands in contrast with much of the economic literature on risk sharing which emphasizes self-interest and reciprocal obligations (Coate and Ravallion 1993, Ligon, Thomas and Worrall 2001).

2 Conceptual framework

We first introduce the conceptual framework that underlies our test of bilateral versus unilateral link formation. We then explain how we correct for the possibility of response bias.

2.1 Bilateral versus unilateral link formation

Consider a risk sharing network. A transfer from households i to household j is denoted τ_{ij}. Because transfers partly respond to shocks affecting i and j, they need not be observed over a fixed time interval even if a sharing link exists between i and j. We also have measures of households' willingness to share risk with the partner, w_{ij} and w_{ji}. These measures are dichotomous, with $w_{ij} = \{0, 1\}$ and $w_{ji} = \{0, 1\}$.

If risk sharing is unilateral, transfers are more likely to take place between i and j whenever either of them wishes to link. In this case the likelihood of observing transfers τ_{ij} increase in both w_{ij} and w_{ji}. If risk sharing is bilateral, transfers will only take place if both i and j wish to link, that is, if $w_{ij}w_{ji} = 1$. Once we control for $w_{ij}w_{ji}$, variables w_{ij} and w_{ji} should have no additional effect on the probability of a transfer.

This suggests the following testing strategy. Estimate a regression model of the form:

$$\tau_{ij} = \lambda(\alpha w_{ij} + \beta w_{ji} + \gamma w_{ij}w_{ji} + \theta X_{ij}) \tag{1}$$

where X_{ij} is a vector of controls and λ is the logit function. If risk sharing is unilateral, the likelihood of transfer is the same whether $\{w_{ij}, w_{ji}\} = \{1, 0\}, \{0, 1\}$, or $\{1, 1\}$. It follows that:

$$\alpha = \beta = \alpha + \beta + \gamma > 0$$

which implies that $\gamma = -\beta = -\alpha$. If risk sharing is bilateral, transfers arise only if $\{w_{ij}, w_{ji}\} = \{1, 1\}$. It follows that $\alpha = \beta = 0$ and $\gamma > 0$. Our objective is to estimate the regression model (1).

An alternative test can be constructed using the reverse of model (1). The usefulness of this alternative test will become clear once we introduce response bias. Let $h_{ij} = 1$ if $\tau_{ij} = 0$, i.e., h_{ij} is an indicator variable that takes value 1 is i does *not* give something to j. Similarly define $u_{ij} = 1 - w_{ij}$, i's *un*willingness to link with j.

In the unilateral model of link formation, $h_{ij} = 1$ if both i and j are unwilling to link, i.e., if $\{u_{ij}, u_{ji}\} = \{1, 1\}$. In contrast, in a bilateral model of link formation, $h_{ij} = 1$ if either i or j are unwilling to link, i.e., if $\{u_{ij}, u_{ji}\} = \{1, 0\}, \{0, 1\}$, or $\{1, 1\}$. Estimate a model of the form:

$$h_{ij} = \lambda(\alpha' u_{ij} + \beta' u_{ji} + \gamma' u_{ij} u_{ji} + \theta' X_{ij}) \tag{2}$$

If risk sharing is unilateral, transfers do not take place only if $\{u_{ij}, u_{ji}\} = \{1, 1\}$. It follows that $\alpha' = \beta' = 0$ and $\gamma' > 0$. In contrast, if risk sharing is bilateral, we have $\gamma' = -\beta' = -\alpha' < 0$.

2.2 Response bias

In our data, both i and j were asked about transfers (i.e. loans and gifts) between them. In principle, i and j should report the same transfers τ_{ij}. This is not, however, what we observe: where one side reports $\tau_{ij} > 0$, the other typically reports $\tau_{ij} = 0$. We have no reason to suspect that respondents report transfers that did not take place.[2] It follows that discrepancies between reports made by i and j must be due to under-reporting.

Dropping the ij subscripts to improve readability, let τ denote the true transfer from i to j, i.e., $\tau = 1$ if i made a transfer to j. Further let G be the report that the giver i made on this transfer and let R be the report that the receiver j made on the same transfer. We have $G = 1$ if i reported making a transfer and 0 otherwise. Similarly, $R = 1$ if j reported receiving a transfer, and 0 otherwise. We do not observe τ, only G and R. We assume no over-reporting, which implies that $G = 1$ only if $\tau = 1$, and that $R = 1$ only if $\tau = 1$.

Given these assumptions, the data generation process takes the following form:

$$\begin{aligned}
\Pr(G=1, R=0) &= \Pr(\tau = 1, G = 1, R = 0) \\
&= \Pr(\tau = 1) * \Pr(G = 1|\tau = 1) * \Pr(R = 0|G = 1, \tau = 1) \\
\Pr(G=0, R=1) &= \Pr(\tau = 1, G = 0, R = 1) \\
&= \Pr(\tau = 1) * \Pr(G = 0|\tau = 1) * \Pr(R = 1|G = 0, \tau = 1) \\
\Pr(G=1, R=1) &= \Pr(\tau = 1, G = 1, R = 1) \\
&= \Pr(\tau = 1) * \Pr(G = 1|\tau = 1) * \Pr(R = 1|G = 1, \tau = 1) \\
\Pr(G=0, R=0) &= 1 - \Pr(G=1, R=0) - \Pr(G=0, R=1) - \Pr(G=1, R=1)
\end{aligned}$$

If we further assume that under-reporting by i is independent of under-reporting by j, $\Pr(R|G, \tau) = \Pr(R|\tau)$. This assumption, which is required for identification, is reasonable if

[2] The main reason is that reporting to an enumerator that a transfer has taken place takes time and effort while not reporting it does not. Wrongly reporting a transfer that did not take place requires effort but does not generate any benefit, given that all enumerators are external to the village.

under-reporting results primarily from reporting mistakes and omissions. With this assumption, we can rewrite the system as:

$$\begin{aligned} \Pr(G=1,R=0) &= \Pr(\tau=1)*\Pr(G=1|\tau=1)*\Pr(R=0|\tau=1) && (3)\\ \Pr(G=0,R=1) &= \Pr(\tau=1)*\Pr(G=0|\tau=1)*\Pr(R=1|\tau=1) && (4)\\ \Pr(G=1,R=1) &= \Pr(\tau=1)*\Pr(G=1|\tau=1)*\Pr(R=1|\tau=1) && (5)\\ \Pr(G=0,R=0) &= 1-\Pr(G=1,R=0)-\Pr(G=0,R=1)-\Pr(G=1,R=1) && (6) \end{aligned}$$

Equations (3) to (6) express the data generating process in terms of three probabilities: $P(\tau=1)$, $P(G=1|\tau=1)$ and $P(R=1|\tau=1)$. We assume that these three probabilities can be represented by three distinct logit functions $\lambda(.)$ as follows:

$$\begin{aligned} \Pr(\tau=1) &= \lambda_T(\beta_\tau X_\tau) && (7)\\ \Pr(G=1|\tau=1) &= \lambda_G(\beta_G X_G) && (8)\\ \Pr(R=1|\tau=1) &= \lambda_R(\beta_R X_R) && (9) \end{aligned}$$

Together with (3) to (6), equations (7) to (9) fully characterize the likelihood of observing the data. The main equation of interest is $\lambda_T(\beta_\tau X_\tau)$ which corresponds to equation (1): it is on this equation that we wish to test the restrictions imposed by our testing strategy. Conditioning on X_G and X_R in $\Pr(G=1|\tau=1)$ and $\Pr(R=1|\tau=1)$ allows for correlation on observables in reporting probabilities between the giving and receiving households.

In the literature to date, response bias has typically been ignored and estimation has proceeded using transfers τ_{ij} reported by i, j, or a combination of the two. For instance, Fafchamps and Lund (2003) and De Weerdt and Fafchamps (2009) use transfers information obtained from one of the two households only – i for transfers given, and j for transfers received. Fafchamps and Gubert (2007) combine answers given by i and j to construct a unique measure of τ_{ij}. Whether or not response bias affects inference probably depends on the hypothesis being tested.

Our ultimate objective is to test whether gifts and loans are unilateral or bilateral. Hence we are primarily interested in the coefficients of w_{ij}, w_{ji} and $w_{ij}w_{ji}$ in equation (7). We expect the propensity to report a gift to vary systematically with w_{ij} and w_{ji}. For instance, i may be more likely to report gifts to households from whom he wishes to seek help in the future, i.e., households for which $w_{ij}=1$. Similarly, j may be more likely to report gifts received from households for which $w_{ji}=1$.

To investigate whether response bias may affect inference regarding the coefficients of w_{ij}, w_{ji} and $w_{ij}w_{ji}$ in equation (7), we conduct a simulation analysis of the data generating process defined by equations (3) to (9) under different assumptions regarding response bias and link formation. Results, not shown here to save space, show that response bias can dramatically affect inference regarding w_{ij}, w_{ji} and $w_{ij}w_{ji}$.

If we observe the actual gifts τ_{ij} without reporting bias, equation (1) can be estimated directly. Results are as anticipated: if link formation is bilateral, $\alpha = \beta = 0$ while $\gamma > 0$; if link formation is unilateral, $\gamma = -\beta = -\alpha$ holds. If we do not observe the actual gifts τ_{ij}, we can choose to ignore reporting bias and estimate equation (1) by assuming that a transfer took place if *either* i or j reported it. Simulation results indicate that, in this case, coefficient estimates (1) are reasonable if the reporting bias does not depend on willingness to link w_{ij} and w_{ji}. However, if it does, they are severely biased.[3]

Next we estimate equation (1) by maximum likelihood using the likelihood function defined by equations (3) to (9). We first assume that response bias is present but does not depend on w_{ij} and w_{ji}. In this case, ML estimates are consistent whether or not we include w_{ij} in X_G and w_{ji} in X_R. We then assume that w_{ij} is in X_G and w_{ji} is in X_R. This is equivalent to assuming that respondents are more likely to remember a transfer to (or from) individuals with whom they wish to share mutual assistance. In this case, ML estimates are consistent only if we include w_{ij} in X_G and w_{ji} in X_R. If we do not, the coefficient of $w_{ij}w_{ji}$ – which is essential to our testing strategy – is severely biased, often with the incorrect sign. Based on these results, we estimate model (3) to (9) with w_{ij} in X_G and w_{ji} in X_R.

Simulations also show that, if response bias does not depend on willingness to link, consistent ML estimates obtain even if X_G and X_R only contain an intercept term. This indicates that identification does not require that X_G and X_R contain a variable that does not enter X_τ.

2.3 Reverse model

In order to estimate model (3) to (9), we had to assume that if a transfer is reported by either i or j, it necessarily took place. This assumption was critical to identify model (3) to (9) from the observed data. We want to know whether our inference regarding unilateral versus bilateral link formation is sensitive to this assumption. In particular, we want to investigate whether test results change when we assume that all discordant answers are due to over-reporting, i.e., to people reporting transfers that did not actually take place. This could arise, for instance, because people wish they had made these transfers but were ashamed to admit this to enumerators, and so made up some numbers. Whether or not this is a reasonable assumption depends on the context. In the context of our data, we feel that this is unlikely. But we still wish to investigate the robustness of our results to this assumption.

[3] We also estimated equation (1) assuming that a transfer took place if *both* i and j report it. In this case, results are inconsistent irrespective of the form of response bias. This is hardly surprising given the assumed data generation process precludes over-reporting.

Formally, we now wish to assume that unless both i and j declare a transfer, it did not take place. As before, we assume that response errors are independent between i and j, an assumption that is required for identification. We have:

$$\begin{aligned}
\Pr(G=1,R=0) &= \Pr(\tau=0)*\Pr(G=1|\tau=0)*\Pr(R=0|\tau=0) &(10)\\
\Pr(G=0,R=1) &= \Pr(\tau=0)*\Pr(G=0|\tau=0)*\Pr(R=1|\tau=0) &(11)\\
\Pr(G=0,R=0) &= \Pr(\tau=0)*\Pr(G=0|\tau=0)*\Pr(R=0|\tau=0) &(12)\\
\Pr(G=1,R=1) &= 1-\Pr(G=1,R=0)-\Pr(G=0,R=1)-\Pr(G=0,R=0) &(13)
\end{aligned}$$

Equations (3) to (6) express the data generating process in terms of three probabilities: $P(\tau=0)$, $P(G=1|\tau=0)$ and $P(R=1|\tau=0)$. As before, we assume that these three probabilities can be represented by three distinct logit functions $\lambda(.)$ as follows:

$$\begin{aligned}
\Pr(\tau=0) &= \lambda_T(\beta_\tau' X_\tau) &(14)\\
\Pr(G=1|\tau=0) &= \lambda_G(\beta_G' X_G) &(15)\\
\Pr(R=1|\tau=0) &= \lambda_R(\beta_R' X_R) &(16)
\end{aligned}$$

The main equation of interest now is $\Pr(\tau=0)$. Our objective remains to test whether transfers are unilateral or bilateral. So we need a testing strategy in terms of $\tau=0$. Such a strategy is provided by model (2) with $\Pr(\tau=0)=\Pr(h=1)$.

2.4 Standard errors

Dyadic observations such as those on τ_{ij} (or h_{ij}) are typically not independent. This does not invalidate the application of standard maximum likelihood techniques to estimate β_τ, β_G and β_R. But standard errors must be adjusted to correct for dyadic dependence across observations, otherwise inference will be inconsistent.

If we had data from a sufficient number of distinct sub-populations we could cluster of standard errors to correct for correlation across observations from the same sub-population (Arcand and Fafchamps 2008). Unfortunately, we only have data from a single village. Given this, we apply the formula developed by Fafchamps and Gubert (2007), using the scores in lieu of X. This approach corrects for arbitrary correlation across all τ_{ij} and τ_{ji} observations involving either i or j.

The simulation analysis reported earlier was conducted using dyadic standard errors. Results indicate that t-values obtained via this method are a good basis for correct inference about α, β and γ.

3 The data

We use survey data from Nyakatoke, a village community in the Buboka Rural District of Tanzania, at the west of Lake Victoria. The village's main livelihood is the farming of bananas, sweet potatoes and cassava for food, while coffee is the main cash crop. The community is composed by 600 inhabitants, 307 of which are adults, for a total of 119 households interviewed in five regular intervals during 2000. This dataset is ideal for our purpose because it is a census covering all 119 households in the village.[4] The data include information on households' demographics (composition, age, religion, education), wealth and assets (land and livestock ownership, quality of housing and durable goods), income sources and income shocks, transfers and network relations.

In each of the 5 rounds of data collection (February to December 2000) each adult household member were asked whether they had received or given any loans or gifts.[5] If they said yes, information was collected on the name of the partner, the value of what was given or received, whether in cash or kind. Loan repayment and gifts in labour are not included. Aggregating at the household level across rounds, we obtain a picture of transfers of funds between all households in the village. Aggregating across rounds should reduce discrepancies in answers due to difference in interview dates across households. Aggregating at the household level should also reduce discrepancies, e.g., if i mentioned giving to member a of household j but member b is the one who mentions receiving a gift from household i.

For each household dyad ij we have four variables: gifts τ_{ij}^i that i stated giving to j; gifts τ_{ij}^j that j stated receiving from i; gifts τ_{ji}^j that j declared giving to i; and gifts τ_{ji}^i that i stated receiving from j. Similar data is available for loans. The literature on informal risk sharing has noted that informal loans often serve to smooth consumption against shocks (Udry 1994) and can be a way of reducing self-enforcement constraints (Foster and Rosenzweig 2001, Kocherlakota 1996, Ligon and Thomas and Worrall 2001, Fafchamps 1999). In Nyakatoke, gifts are more frequent than loans but smaller in magnitude (De Weerdt and Dercon 2006, De Weerdt and Fafchamps 2009). This is in line with findings reported by Fafchamps and Lund (2003) for the Philippines. Gifts in Nyakatoke have been shown to serve an insurance purpose against health shocks (De Weerdt and Fafchamps, 2009).

There are major discrepancies between τ_{ij}^i and τ_{ij}^j. In fact, $\tau_{ij}^i \neq \tau_{ij}^j$ in nearly all cases, especially for loans. There are 1420 dyads (i.e., 10% of the dyads) for which τ_{ij}^i or τ_{ij}^j is not zero for gifts. Of those, in 42% of cases the report comes from the giver only, in 30% from the receiver only, and in 27% from both. For inter-household loans, there are 545 dyads (i.e.,

[4]Everyone in the village agreed to participate in the survey, but there are some missing data for 4 households.

[5]We do not have loan and gift information at the individual level. When aggregating at the household level, questionnaires were carefully checked by survey supervisors to avoid any double-counting of identical gifts reported by two different members of the same household.

4% of the dyads) for which either i or j reports a loan from i to j. In 56% of these cases, the report comes from the giver only, in 36% from the receiver only, and in 8% from both. Out of 378 dyads in which both i and j report a gift from i to j, only 22 report the same amount. For loans, the corresponding number is 5 out of 37. When the amounts declared differ, they differ by a large margin: the highest of the two declared amounts is on average double the smallest one. This is true for both loans and gifts. The frequency distribution of loan and gift amounts is given in Table A1 in Appendix.

We also checked whether discrepancies may be due to the fact that respondents mix up loans and gifts. The within-dyad correlation between the difference in reported loans and the difference in reported gifts is indeed negative, as would be the case if, say, i reports giving a loan while j reports receiving a gift. But the correlation is small and not statistically significant: if we restrict the sample to the dyads for which at least one loan or gift was declared, the correlation between the difference in reported loans and the difference in reported gifts is -0.036 with a significance level of 0.13.

To summarize, there are massive discrepancies between the responses given by i and j about the same gifts and loans τ_{ij}. These discrepancies are mostly due to the fact that in the the large majority of cases – 93% of the cases for loans and 73% of the cases for gifts – one side reports something while the other reports nothing. Under-reporting by those who receive gifts and loans may not be too surprising: they may have a strategic motive in 'forgetting' the favors they probably have a moral obligation to reciprocate. But we also observe massive under-reporting by those who give. Consequently there may be many transfers which took place but are not observed in the data because they were not mentioned by either sides. When estimating model (2), our main challenge is to address this bias.

3.1 Variables definition

Our unit of observation is the dyad: in Nyakatoke there are 119 households, which gives $119 * 118 = 14042$ possible dyads. We organize the data such that the first listed household refers to the giver and the second to the receiver, i.e., τ_{ij} refers to a transfer from i to j. Note that τ_{ij} defines a directed network: τ_{ij} represents the transfer from i to j, while τ_{ji} represents the transfer from j to i. For τ_{ij} we have two different measurements: the information provided by the giver τ_{ij}^i, and the information provided by the receiver τ_{ij}^j. Similarly for τ_{ji}.

From equation (1) our main regressors of interest are w_{ij}, w_{ji} and $w_{ij}w_{ji}$. In the first Nyakatoke survey round (February 2000), each adult household member was asked: *"Can you give a list of people from inside or outside of Nyakatoke, who you can personally rely on for help and/or that can rely on you for help in cash, kind or labor?"* Answers to this question, aggregated at the household level, are used a proxies for w_{ij} and w_{ji}. This requires some explanation given that the question in principle asks about existing links – not willingness to link. We first note that if responses perfectly captured actual links, then we would observe $w_{ij} = w_{ji}$ for all i and j. This is not the case: out of 14042 possible dyads, there are 980 dyads for which w_{ij} or w_{ji} is not 0. Of those, only 280 have $w_{ij} = w_{ji} = 1$ while 700 dyads have $w_{ij} = 1$ but $w_{ji} = 0$ or the reverse.

There remains the possibility that w_{ij} and w_{ji} are about actual links but contain a lot of mis-reporting. Comola and Fafchamps (2009) examine this issue in detail using the same data. They test whether w_{ij} and w_{ji} are best viewed as willingness to link or as misreported links. They find that the data are best interpreted as willingness to link. Identification is achieved by noting that, if w_{ij} and w_{ji} measure willingness to link, i lists nodes j that are attractive to i irrespective of whether i is attractive to j. In contrast, if w_{ij} and w_{ji} are two statements about the same actual link, i should take into account i's own attractiveness to j when answering the question.

We implement a simplified version of their test as follows. Let z_j be a characteristic of j correlated with i's willingness to link with j, and similarly for z_i. Stack observations on w_{ij} and w_{ji} and regress them on z_i and z_j in a probit regression of the form $w_{ij} = az_i + bz_j + u_{ij}$ and $w_{ji} = az_j + bz_i + u_{ji}$. Consider what happens if w_{ij} and w_{ji} are measurements of actual links and link formation is bilateral, but i and j sometimes forget to report existing links.[6] In this case, w_{ij} should be 1 only when i knows j wants to link with him. Similarly, w_{ji} should only be 1 when j knows that i wants to link with him. Since both w_{ij} and w_{ji} enter the regression, on average we should have $a \approx b$. A similar prediction arises when link formation is unilateral: i should report a link whenever i OR j wishes to link – and thus the likelihood of reporting a link w_{ij} should rise with both the attractiveness of i and that of j. It is also conceivable that i only mentions those links that he cares about, and j likewise. When this happens, w_{ij} is increasing in the attractiveness of j for i, but not in i's attractiveness to j, i.e., $b > 0$ but $a = 0$. It this case, w_{ij} proxies for i's willingness to link with j, not for a link between i and j.

As predictors of attractiveness z_j we use wealth and popularity: wealthier households are in a better position to assist others in need; and popularity proxies for other attributes correlated with attractiveness. Regression results are reported in Appendix Table A2.[7] We find $b > 0$ – the wealth and popularity of the partner are strong predictors of w_{ij} – but $a = 0$ – own characteristics are not significant. These results confirm that w_{ij} and w_{ji} can reasonably be regarded as proxying for willingness to link.

Turning to other regressors, the main regression of interest $\Pr(\tau = 1) = \lambda_T(\beta_\tau X_\tau)$ seeks to

[6] A similar reasoning applies if respondents over-report links, i.e., report links that do not exist but that they wish existed.

[7] Wealth is computed as the total value of land and livestock assets in Tanzanian shilling (1 unit = 100000 *tzs*). Popularity of household j is defined as the number of times j is listed by households other than i in response to the risk sharing question.

explain the transfers that are made. The regressors entering X_τ are control variables expected to influence the actual flows of funds between households. Since τ_{ij} is directional, regressors for observation ij can differ from regressors for observation ji; this stands in contrast with undirected network data where regressors by construction have to be identical. We expect flows of funds between households to depend on the wealth of the giver and receiver, which we control for. From the work of Fafchamps and Lund (2003), De Weerdt and Dercon (2006) and De Weerdt and Fafchamps (2009), we suspect risk sharing to be more frequent among households that are geographically and socially proximate. Finally, larger households have more individuals involved in giving and receiving transfers. We therefore control for the wealth of i and j, the number of adult members for i and j, the distance between the two houses, and dummies for whether i and j are related, and share the same religion.

Next we discuss the variables that enter $\Pr(G = 1|\tau = 1) = \lambda_G(\beta_G X_G)$ and $\Pr(R = 1|\tau = 1) = \lambda_R(\beta_R X_R)$. The first measures the propensity for the giver to report a transfer that has taken place; the second measures the receiver's propensity to report a transfer that has taken place. As discussed earlier, we include w_{ij} in X_G – givers are more likely to remember transfers to individuals whose name they listed in response to first-round interviews. We include w_{ji} in X_R for the same reason. We also include as regressors the own wealth (wealth of i in X_G and wealth of j in X_G) as wealthy people are more likely to forget a transfer. Social and geographical proximity variables are also included because respondents are more likely to remember transfers to and from proximate households.

We also include regressors that can be expected to affect response bias but not transfers themselves.[8] For X_G, we use $n_i \equiv \sum_j w_{ij}$, that is, the number of individuals listed in response to the first-round question on who respondents would turn for help – to whom they would provide help. The logic underneath this choice is that households intending to seek help from (or provide help to) many other households may be more sensitive to the issue, and therefore recall transfers better. For X_R we include the number of male and female adult dependents. The idea is that adult dependents who have received transfers from other households may not have reported them to the household head – and therefore may be reluctant to report them to enumerators.

To illustrate whether our correction for mis-reporting affects inference regarding the link formation process, we also estimate two probit regressions which are by construction directly comparable with $\Pr(\tau = 1)$. In the first of them, the dependent variable equals one if at least one side has declared a gift. This is equivalent to defining $\tau_{ij}^u \equiv \max\{\tau_{ij}^i, \tau_{ij}^j\}$. This assumes that mis-reporting only takes the form of under-reporting. In the second regression the dependent variable equals one if both the giver and the receiver have declared a gift, i.e., it is $\tau_{ij}^o \equiv \min\{\tau_{ij}^i, \tau_{ij}^j\}$, which is equivalent to assuming that mis-reporting takes the form of

[8]Simulation analysis reported earlier indicates that ML estimates are reliable even without identifying instruments, so including these variables is not necessary for identification.

over-reporting.

In Table 1 we present summary statistics for all variables used in the analysis. The upper section of the table reports different 'versions' of the dependent variable. The first two rows focus on the gifts from i to j, as reported by i and j. Variables τ_{ij}^i takes value 1 if i reported a gift to j, and 0 otherwise. Similarly for τ_{ij}^j. We see that givers are more likely to report a gift than receivers. In the next two rows we report $\tau_{ij}^u \equiv \max\{\tau_{ij}^i, \tau_{ij}^j\}$ and $\tau_{ij}^o \equiv \min\{\tau_{ij}^i, \tau_{ij}^j\}$. They demonstrate the extent of the divergence between the information given by households i and j on the same reality τ_{ij}.

In the next four rows we report the same information for inter-household loans. Variables are constructed in the same way. Here too we see that lenders are more likely to report a loan than borrowers, and that there are considerable discrepancies between loans reported given and loans reported received.

From these figures it is possible to compute a rough estimate of extent of under-reporting, assuming independence in reporting probability between i and j. We focus on gifts first. We wish to estimate three unconditional probabilities: $\Pr(\tau = 1)$, $\Pr(G = 1|\tau = 1)$, and $\Pr(R = 0|\tau = 1)$. We have three equations to do so:

$$\Pr(G = 1, R = 0) = \Pr(\tau = 1) * \Pr(G = 1|\tau = 1) * \Pr(R = 0|\tau = 1) = 0.043 \quad (17)$$
$$\Pr(G = 0, R = 1) = \Pr(\tau = 1) * \Pr(G = 0|\tau = 1) * \Pr(R = 1|\tau = 1) = 0.031 \quad (18)$$
$$\Pr(G = 1, R = 1) = \Pr(\tau = 1) * \Pr(G = 1|\tau = 1) * \Pr(R = 1|\tau = 1) = 0.028 \quad (19)$$

Simple algebra yields the following solution:

$$\begin{aligned} \Pr(\tau = 1) &= 15\% \\ \Pr(G = 1|\tau = 1) &= 47\% \\ \Pr(R = 1|\tau = 1) &= 39\% \end{aligned}$$

The above calculation shows that there is considerable under-reporting of gifts and that $\tau_{ij}^u = 10.1\%$ underestimates the frequency of gifts by almost 50%. A similar calculation for loans yields:

$$\begin{aligned} \Pr(\tau = 1) &= 14\% \\ \Pr(G = 1|\tau = 1) &= 18\% \\ \Pr(R = 1|\tau = 1) &= 12\% \end{aligned}$$

which suggests massive under-reporting of loans and indicates that $\tau_{ij}^u = 3.9\%$ only captures a quarter of the loans we suspect were made.

Table 1. Descriptive statistics (N=14042)

variable	mean	min	max	sd
τ_{ij}^i (gifts)	0.071			
τ_{ij}^j (gifts)	0.059			
τ_{ij}^u (gifts)	0.101			
τ_{ij}^o (gifts)	0.028			
τ_{ij}^i (loans)	0.025			
τ_{ij}^j (loans)	0.017			
τ_{ij}^u (loans)	0.039			
τ_{ij}^o (loans)	0.003			
w_{ij} and w_{ji}	0.045			
$w_{ij}w_{ji}$	0.020			
weighted w_{ij} and *weighted* w_{ji}	0.023	0	0.933	0.117
$wealth_i$ and $wealth_j$	4.546	0	27.970	4.815
same religion	0.354			
related	0.016			
distance	0.522	0.014	1.738	0.303
$hh\,members_i$ and $hh\,members_j$	2.555	1	9	1.314
n_i	5.294	0	19	3.063
$female\ dependents_j$	1.101	0	6	0.864
$male\ dependents_j$	0.437	0	3	0.729

The rest of the table focuses on regressors. Variable $w_{ij} = 1$ if someone in household i mentioned someone in household j in response to the first-round question on who respondents turn for help. The product $w_{ij}w_{ji} = 1$ if i listed j and j listed i, something that occurs only for 2% of the dyads.

We also report a weighted version of w_{ij} that is constructed as follows. Remember that the first-round question on who respondents turn for help was answered separately by each adult members of the household. For each household member l in household i, we know the order in which they listed various individuals m from other households j. This order may contain information on how seriously l regards m to be a possible source of assistance. To aggregate this information at the level of the household, we construct a weighted link variable $weight_{lm}$ for each lm pair. This variable is defined as:

$$weight_{lm} = \frac{(\#names_l + 1) - rank_{lm}}{\#names_l + 1}$$

where $\#names_l$ is the total number of names given by l and $rank_{lm}$ is the order in which m was listed by l. We then average $weight_{lm}$ across all l members of household i and all m

members of household j.[9]

Control variables are reported next. Whenever the average is the same for giver and receiver, we only report one of them. Wealth is computed as the total value of land and livestock assets (1 unit = 100000 tzs). We see there is considerable variation in wealth levels across Nyakatoke households. There is also significant diversity in religion: only 35% of households heads share the same religion.[10] Around 1.6% of household pairs are closely related, i.e., are siblings or children-parents. Distance between households is measured in Km and is on average 500 meters.[11] Adult members are those aged 15 and above. Male and female dependents are defined as adult members of the household who are not the head of household. Wives are included in the dependents, the idea being that they too may seek to dissimulate gifts and loans received from other households.

4 Estimation results

4.1 Baseline model

We now proceed to estimate equation (1). In the results presented in Table 2 transfers τ_{ij} refer to gifts from i to j, in cash or in kind. Columns (1) and (2) report simple probit regressions where the dependent variable is τ_{ij}^u and τ_{ij}^o, respectively. These regressions are presented as robustness checks on estimates presented in column (3). Columns (3) to (5) of the table report coefficients obtained from estimating the likelihood function formed by combining equations (3) to (9). Column (3) corresponds to our model of interest (1).

Results strongly reject the bilateral link formation model: both α and β are strongly significant, while γ is never significantly positive. Coefficient estimates are at least partly consistent with unilateral link formation: α and β, the coefficients of w_{ij} and w_{ji}, are both significant and of the same order of magnitude. A Wald test cannot reject the hypothesis that $\alpha = \beta$ in column (3), with a p-value of 0.3652. This is true in the model estimates that correct for under-reporting, but also in columns (1) and (2) where the dependent variable is constructed in a more conventional manner. As predicted by the unilateral model the coefficient γ of $w_{ij}w_{ji}$ is negative in all three columns (1), (2) and (3) – but only significantly so in column (1). Furthermore, contrary to the predictions of the unilateral link formation model, $\gamma \neq -\beta$ and $\gamma \neq -\alpha$: a Wald test rejects the joint hypotheses $\gamma + \alpha = 0$ and $\gamma + \beta = 0$ with p-value=0.002. This means that if $w_{ij} = 1$ and also $w_{ji} = 1$ then the probability of transfer is larger than if either of them alone is equal to 1. In other words, when both

[9] Whenever l mentions someone who lives outside Nyakatoke, we take this person into account when computing $\#names_l$ and $rank_m$.

[10] Out of 119 households, 24 are Muslim (20%), 46 are Protestant (39%) and 49 are Catholic (41%).

[11] For 3 households the distance to other households is missing, so we have imputed the sample average to avoid losing those observations.

households list each other as someone they would go to for help, they are more likely to help each other than if only one lists the other. This suggests that some bilateral dimension is present, even if the results strongly reject the bilateral model itself.

Control variables have reasonable coefficients in columns (1), (2) and (3). Wealthier households give more (and may also receive more – see column 1). People give more to relatives, neighbors, and members of the same religious community.

Results for the two under-reporting regressions – columns (4) and (5) – show that respondents are more likely to report a transfer from/to those households they have previously mentioned.

In the $\Pr(G = 1|\tau = 1)$ regression, w_{ij} is positively significant, indicating that if household i has listed household j in the sense that $w_{ij} = 1$, then i is more likely to report a gift given to j. Variable n_i, which is the total number of individuals listed in response of the first-round question, is significantly positive, suggesting that large households are more likely to report gifts given. Also, own wealth is significant and negative: wealthy respondents are more likely to forget reporting the gifts they have made. Analogously, in the $\Pr(R = 1|\tau = 1)$ regression w_{ji} is positively significant, and $wealth_j$ is negatively significant. The numbers of female and male dependents have the anticipated negative sign, but they are not significant.

To get a sense of the magnitude of the under-reporting effects, we calculate marginal effects for the $\Pr(G = 1|\tau = 1)$ and $\Pr(R = 1|\tau = 1)$ regressions. Results, reported in Table A3 in appendix, confirm that w_{ij} and w_{ji} have quantitatively the largest effect on reporting bias. Relatedness and geographical distance also have effects that are large in magnitude.

In Table 3 we repeat the same analysis using loans instead of gifts. Coefficient estimates reported in columns (1) and (3) more or less satisfy $\alpha = \beta = -\gamma$, a finding that is consistent with the unilateral model. A Wald test of the joint hypothesis that $\alpha = \beta = -\gamma$ has p-value of 0.9303, implying that we cannot reject the hypothesis of unilateral link formation. But these coefficients are only statistically significant in column (1). This may be because the proportion of non-zero observations is smaller for loans, making ML estimation more difficult. In terms of the other regressors, few of them are significant, a point already noted by De Weerdt and Fafchamps (2009) in the same dataset. In column (3), we find $wealth_i$ (marginally) significant, reconfirming the intuition that wealthy households are those who lend money. In the $\Pr(G = 1|\tau = 1)$ regression only n_i is significantly positive, and in the $\Pr(R = 1|\tau = 1)$ regression only the previously declared willingness to link w_{ji} is significantly positive. Marginal effects reported in Table A2 show that the variables with the largest impact are w_{ij} (which is not significant) and n_i for the giver and w_{ji} for the receiver.

Table 2. Results for gifts

	(1) τ_{ij}^u	(2) τ_{ij}^o	(3) $\Pr(\tau = 1)$	(4) $\Pr(G = 1 \vert \tau = 1)$	(5) $Pr(R = 1 \vert \tau = 1)$
w_{ij}	1.401***	1.129***	2.563***	1.492***	
	(0.107)	(0.116)	(0.371)	(0.180)	
w_{ji}	1.582***	1.521***	2.817***		1.920***
	(0.093)	(0.109)	(0.305)		(0.227)
$w_{ij}w_{ji}$	-0.417**	-0.235	-0.196		
	(0.194)	(0.165)	(0.980)		
$wealth_i$	0.029***	0.013*	0.081***	-0.035**	
	(0.007)	(0.008)	(0.016)	(0.016)	
$wealth_j$	0.035**	0.000	0.105		-0.045***
	(0.018)	(0.012)	(0.066)		(0.015)
$same\,religion$	0.221***	0.174**	0.530**	0.025	0.012
	(0.052)	(0.068)	(0.251)	(0.211)	(0.196)
$related$	0.942***	0.583***	1.961**	0.433	0.614
	(0.173)	(0.189)	(0.762)	(0.505)	(0.377)
$distance$	-0.829***	-0.892***	-1.678**	-0.585	-0.533
	(0.186)	(0.317)	(0.660)	(0.536)	(0.485)
$hh\,members_i$	0.047***	0.038	0.110**		
	(0.018)	(0.028)	(0.043)		
$hh\,members_j$	0.108**	0.082*	0.262		
	(0.054)	(0.044)	(0.168)		
n_i				0.026*	
				(0.013)	
$female\ dependents_j$					-0.149
					(0.143)
$male\ dependents_j$					-0.191
					(0.133)
$constant$	-2.016***	-2.511***	-3.525***	-0.277	-0.209
	(0.156)	(0.191)	(0.540)	(0.590)	(0.359)

*** p<0.01, ** p<0.05, * p<0.1. Dyadic-robust standard errors in parentheses.

Table 3. Results for loans

	(1) τ_{ij}^{u}	(2) τ_{ij}^{o}	(3) $\Pr(\tau=1)$	(4) $\Pr(G=1\vert\tau=1)$	(5) $Pr(R=1\vert\tau=1)$
w_{ij}	0.969***	0.464	2.639	0.570	
	(0.112)	(0.413)	(5.599)	(0.624)	
w_{ji}	1.006***	1.169***	2.536		1.206**
	(0.106)	(0.191)	(6.437)		(0.558)
$w_{ij}w_{ji}$	-0.750***	-0.195	-2.021		
	(0.116)	(0.421)	(8.388)		
$wealth_i$	0.010	-0.005	0.061*	-0.041	
	(0.009)	(0.013)	(0.036)	(0.051)	
$wealth_j$	0.008	0.006	0.031		-0.012
	(0.006)	(0.006)	(0.051)		(0.031)
$same\,religion$	0.080	-0.082	0.323	-0.058	-0.041
	(0.055)	(0.219)	(2.717)	(1.601)	(1.048)
$related$	0.107	-0.100	0.681	-0.079	0.133
	(0.170)	(0.376)	(18.080)	(1.946)	(1.760)
$distance$	-0.576***	-0.471**	-1.775	-0.083	0.020
	(0.125)	(0.231)	(1.282)	(1.608)	(1.191)
$hh\,members_i$	0.023	0.028	0.013		
	(0.040)	(0.048)	(0.270)		
$hh\,members_j$	0.021	0.018	0.192		
	(0.026)	(0.055)	(0.635)		
n_i				0.113**	
				(0.047)	
$female\ dependents_j$					-0.047
					(0.168)
$male\ dependents_j$					-0.222
					(0.157)
$constant$	-1.915***	-2.969***	-1.991	-2.478	-2.442*
	(0.136)	(0.223)	(2.032)	(2.208)	(1.409)

*** p<0.01, ** p<0.05, * p<0.1. Dyadic-robust standard errors in parentheses.

4.2 Reverse model

Next we estimate the reverse model, which assumes that all discordant answers are due to over-reporting. The testing strategy and likelihood function corresponding to this assumption were presented in Section 2. It should be noted that, in our data, only 10% of household pairs both declare a gift and only 3% both declare a loan. This means that, under the assumption of no under-reporting, the number of loan observations for which $\tau = 1$ is small, making inference more difficult and possibly creating identification and convergence problems. It is nevertheless instructive to investigate whether we obtain results that do not contradict our earlier conclusions regarding unilateral link formation.

Estimation results are presented in Table 4 for gifts and Table 5 for loans. Results are less conclusive than those reported in Tables 2 and 3. Coefficients α' and β' are significantly positive in all three gift regressions (Table 4), that is, for the two logit models and for the ML model that corrects for response bias. In the loan regressions (Table 5), α' and β' are positive in all three regressions, although only significantly so in the logit regressions. As explained in section 2 when discussing equation (2), this evidence is consistent with the unilateral link formation hypothesis. However, γ', the coefficient of $(1 - w_{ij})(1 - w_{ji})$, is also positive and significant in several of the regressions, which is consistent with bilateral link formation. Hence, when we assume that there is no under-reporting (only over-reporting), the evidence is ambiguous in the sense that it supports both models – or a hybrid of the two, where links are formed in a way that is largely unilateral but contains some bilateral element as well.

Regarding the reporting equations, we find, as before, that the likelihood of reporting a gift increases in w_{ij} and w_{ji}. It also increases significantly with kinship, geographical proximity, and co-religion. For loans (Table 5), results show that the likelihood of reporting a loan increases with w_{ij} and w_{ij} and with geographical proximity.

4.3 Robustness analysis

To further investigate the robustness of our results, we reestimate ML model (3) to (6) omitting w_{ij} and w_{ji} from the response bias equations. Results, not shown here to save space, are dramatically different from those reported in Tables 2 and 3. In particular, the coefficient of $w_{ij}w_{ji}$ in the $\Pr(\tau = 1)$ equation becomes large and positive, and has a large t-value. Virtually identical results for $\Pr(\tau = 1)$ are obtained if the $\Pr(G = 1|\tau = 1)$ and $Pr(R = 1|\tau = 1)$ only include an intercept (see Table A4 in Appendix). On the other side, if we only include w_{ij} and w_{ji} and the constant in the response bias equations, results are very close to those reported in Tables 2 and 3. These findings are consistent with the simulation results reported in Section 2. They confirm that correct inference requires that we include w_{ij} and w_{ji} in the response bias equations, as done in Tables 2 and 3.

We worry that what household i reported as a gift was reported as a loan by j. Misclassification would affect estimated reporting propensities and hence may affect inference. To investigate whether misclassification may have affected our results, we reestimate the baseline model using combined gifts and loans as the dependent variable. Results, shown in Table A5 in Appendix, are very similar to those reported in Table 2 for gifts. Misclassification therefore does not seem to explain our results. In the reporting equation for transfer recipients, the number of male dependents is negative and significant at the 10% level. This provides some support to the idea that under-reporting of gifts received is to avoid detection by other

Table 4. Results for gifts, reverse model

	(1) τ^u_{ij}	(2) τ^o_{ij}	(3) $\Pr(\tau = 0)$		(4) $\Pr(G = 1 \mid \tau = 0)$	(5) $Pr(R = 1 \mid \tau = 0)$
$1 - w_{ij}$	0.984***	0.894***	1.143***	w_{ij}	2.724***	
	(0.175)	(0.118)	(0.323)		(0.251)	
$1 - w_{ji}$	1.164***	1.286***	1.889***	w_{ji}		3.036***
	(0.167)	(0.094)	(0.209)			(0.249)
$(1 - w_{ij})(1 - w_{ji})$	0.417**	0.235	1.595***	$wealth_i$	0.026***	
	(0.194)	(0.165)	(0.508)		(0.009)	
$wealth_i$	-0.029***	-0.013*	-0.022	$wealth_j$		0.029
	(0.007)	(0.008)	(0.020)			(0.019)
$wealth_j$	-0.035**	0.000	0.007	$same\,religion$	0.353***	0.371***
	(0.018)	(0.012)	(0.027)		(0.107)	(0.132)
$same\,religion$	-0.221***	-0.174**	-0.286	$related$	1.417	1.914***
	(0.052)	(0.068)	(0.201)		(0.976)	(0.556)
$related$	-0.942***	-0.583***	0.048	$distance$	-1.200**	-1.208**
	(0.173)	(0.189)	(1.477)		(0.511)	(0.508)
$distance$	0.829***	0.892***	1.601	n_i	0.026*	
	(0.186)	(0.317)	(1.110)		(0.015)	
$hh\,members_i$	-0.047***	-0.038	-0.068	$female\ dependents_j$		-0.172*
	(0.018)	(0.028)	(0.074)			(0.101)
$hh\,members_j$	-0.108**	-0.082*	-0.214**	$male\ dependents_j$		-0.071
	(0.054)	(0.044)	(0.107)			(0.148)
$constant$	-0.550**	0.097	0.686**	$constant$	-3.170***	-3.246***
	(0.226)	(0.134)	(0.285)		(0.322)	(0.288)

*** p<0.01, ** p<0.05, * p<0.1. Dyadic-robust standard errors in parentheses.

Table 5. Results for loans, reverse model

	(1) τ_{ij}^u	(2) τ_{ij}^o	(3) Pr($\tau = 0$)		(4) Pr($G = 1\|\tau = 0$)	(5) $Pr(R = 1\|\tau = 0)$
$1 - w_{ij}$	0.219*	0.268*	0.023	w_{ij}	1.881***	
	(0.114)	(0.141)	(0.572)		(0.323)	
$1 - w_{ji}$	0.256**	0.973**	3.037	w_{ji}		2.355***
	(0.111)	(0.469)	(2.577)			(0.321)
$(1 - w_{ij})(1 - w_{ji})$	0.750***	0.195	0.768	$wealth_i$	0.003	
	(0.116)	(0.421)	(2.541)		(0.011)	
$wealth_i$	-0.010	0.005	0.133*	$wealth_j$		0.007
	(0.009)	(0.013)	(0.079)			(0.023)
$wealth_j$	-0.008	-0.006	-0.018	$same\,religion$	0.192	0.222
	(0.006)	(0.006)	(0.031)		(0.158)	(0.181)
$same\,religion$	-0.080	0.082	0.266	$related$	0.255	0.547
	(0.055)	(0.219)	(0.762)		(0.415)	(0.482)
$related$	-0.107	0.100	12.877***	$distance$	-1.257***	-1.027**
	(0.170)	(0.376)	(0.759)		(0.456)	(0.457)
$distance$	0.576***	0.471**	1.129	n_i	0.114***	
	(0.125)	(0.231)	(1.269)		(0.016)	
$hh\,members_i$	-0.023	-0.028	0.003	$female\ dependents_j$		0.030
	(0.040)	(0.048)	(0.207)			(0.165)
$hh\,members_j$	-0.021	-0.018	-0.107	$male\ dependents_j$		-0.114
	(0.026)	(0.055)	(0.256)			(0.164)
$constant$	0.690***	1.532***	2.615*	$constant$	-4.229***	-4.230***
	(0.153)	(0.331)	(1.335)		(0.275)	(0.315)

*** p<0.01, ** p<0.05, * p<0.1. Dyadic-robust standard errors in parentheses.

household members – a point already made by Anderson and Baland (2002) in their study of spouses' saving behavior.

Next, we repeat the analysis adding the weighted version of w_{ij} and w_{ji} as additional regressors in the reporting equations. Everything else is unchanged. Results, reported in Table 6 for gifts and Table 7 for loans, are very similar to those reported in Tables 2 and 3, and the new variables are not significant, with the exception of weighted w_{ij} for gifts.

We have also re-estimated the model with different sets of regressors. Convergence is generally smooth for a moderately sized set of regressors as the ones of Table 2 and 3, and estimated coefficients for the key regressors (as self-declared links, and relational attributes) are similar across specifications. A few regressors in columns (4) and (5) are sufficient to get stable estimates for $\Pr(\tau = 1)$. However, identification gets more problematic if we include partner's characteristic in the response bias equations (i.e., j's characteristics in $\Pr(G = 1|\tau = 1)$ and i's characteristics in $\Pr(R = 1|\tau = 1)$). The results presented here should thus be interpreted as based on these exclusion assumptions.

Table 6. Results for gifts with weighted w_{ij}

	(1)	(2)	(3)
	$\Pr(\tau = 1)$	$\Pr(G = 1 \vert \tau = 1)$	$Pr(R = 1 \vert \tau = 1)$
w_{ij}	2.578***	1.081***	
	(0.376)	(0.278)	
w_{ji}	2.817***		1.819***
	(0.306)		(0.324)
$w_{ij}w_{ji}$	-0.218		
	(0.987)		
$wealth_i$	0.081***	-0.035**	
	(0.016)	(0.016)	
$wealth_j$	0.105		-0.044***
	(0.067)		(0.015)
$same\, religion$	0.533**	0.024	0.010
	(0.251)	(0.212)	(0.194)
$related$	2.002***	0.415	0.597
	(0.776)	(0.504)	(0.379)
$distance$	-1.675**	-0.585	-0.537
	(0.668)	(0.547)	(0.490)
$hh\, members_i$	0.109**		
	(0.043)		
$hh\, members_j$	0.261		
	(0.169)		
n_i		0.027**	
		(0.013)	
$weighted\ w_{ij}$		0.796*	
		(0.435)	
$female\ dependents_j$			-0.150
			(0.143)
$male\ dependents_j$			-0.189
			(0.133)
$weighted\ w_{ji}$			0.200
			(0.472)
$constant$	-3.526***	-0.283	-0.209
	(0.542)	(0.597)	(0.360)

*** p<0.01, ** p<0.05, * p<0.1. Dyadic-robust standard errors in parentheses.

Table 7. **Results for loans with weighted w_{ij}**

	(1) $\Pr(\tau = 1)$	(2) $\Pr(G = 1\|\tau = 1)$	(3) $Pr(R = 1\|\tau = 1)$
w_{ij}	2.639	0.483	
	(5.653)	(0.548)	
w_{ji}	2.531		1.260**
	(6.476)		(0.598)
$w_{ij}w_{ji}$	-2.027		
	(8.449)		
$wealth_i$	0.061*	-0.041	
	(0.037)	(0.051)	
$wealth_j$	0.031		-0.012
	(0.052)		(0.031)
$same\,religion$	0.324	-0.058	-0.042
	(2.742)	(1.616)	(1.058)
$related$	0.681	-0.082	0.133
	(18.455)	(1.978)	(1.809)
$distance$	-1.775	-0.081	0.020
	(1.284)	(1.619)	(1.195)
$hh\,members_i$	0.013		
	(0.272)		
$hh\,members_j$	0.192		
	(0.644)		
n_i		0.113**	
		(0.048)	
$weighted\ w_{ij}$		0.171	
		(0.466)	
$female\ dependents_j$			-0.047
			(0.167)
$male\ dependents_j$			-0.222
			(0.158)
$weighted\ w_{ji}$			-0.103
			(0.578)
$constant$	-1.991	-2.479	-2.439*
	(2.038)	(2.220)	(1.418)

*** p<0.01, ** p<0.05, * p<0.1. Dyadic-robust standard errors in parentheses.

4.4 Estimate of the under-reporting bias

A by-product of the estimation of the maximum likelihood model formed by equations (3) to (9) is that we can estimate the extent of under-reporting. This is achieved by comparing the frequency of giving or lending in data to the average frequency of the fitted $\Pr(\tau = 1)$ from Tables 2 (for gifts) and 3 (for loans). The result of these calculations is reported in Table 8.

Table 8. The estimate of the bias

	gifts	loans
average fitted $\Pr(\tau_{ij} = 1)$	0.1568	0.1942
in data: declared by i	0.0709	0.0249
in data: declared by j	0.0587	0.0169
in data: declared by i or j	0.1011	0.0388
average fitted $\Pr(G = 1 \mid \tau = 1)$	0.3742	0.1138
average fitted $\Pr(R = 1 \mid \tau = 1)$	0.3110	0.0729

The average fitted propensity to give from Table 2 is 15%, which is the same figure as the one we obtained in Section 3 without conditioning on regressors. For loans, the average fitted $\Pr(\tau_{ij} = 1)$ of 19% is larger than our earlier estimate of 14%. Based on these results, informal loans between villagers are more frequent than gifts, although much fewer of them are reported in the survey. Comparing these estimates to actually reported gifts and loans, we see that not taking response bias into consideration leads to serious underestimation of the extent of gift giving and, especially, of lending and borrowing between villagers.

Table 8 also reports the average fitted propensity to report giving and receiving respectively. The average propensity for the giver to report a gift is 37% when we condition reporting on individual characteristics, compared to 47% when we do not. For recipients of a gift, the propensity to report is 31%, compared to 39% when we do not condition on individual characteristics. Estimated reporting probabilities are much lower for loans. Lenders are estimated to report only 11% of loans – compared to 18% when we do not condition. Borrowers estimated to report as little as 7% of loans, versus 12% if we do not condition on household characteristics. If anything, estimated propensities to report gifts and loans fall when we allow them to depend on household characteristics.

The Nyakatoke data were collected with an unusually high level of care, using multiple survey rounds and interviewing each household member separately. Yet results suggests massive under-reporting. This casts some doubt on the reliability of reported gifts and loans in household survey. For instance, many studies have found that reported gifts and loans are insufficient to insulate households against shocks. But if actual gifts and loans are much larger, these findings might be called into question.

For comparison purposes, we also report in Table 9 estimated propensities from a model in which response equations do not include willingness to link w_{ij} as regressor. The purpose of this calculation is to investigate the extent to which our results are sensitive to model specification. We find that, if we omit w_{ij} and w_{ji} from the response equations, the predicted response rate increases for both gifts and loans, and consequently the estimated average propensity to give or lend decreases. But the magnitudes remain similar and the results again suggest that there is much more under-reporting in loans than in gifts.

Table 9. No w_{ij} and w_{ij} in response equations

	gifts	loans
average fitted $\Pr(\tau_{ij} = 1)$	0.1473	0.1763
average fitted $\Pr(G = 1 \mid \tau = 1)$	0.4608	0.1346
average fitted $\Pr(R = 1 \mid \tau = 1)$	0.4130	0.0943

5 Discussion

We have shown that gifts and informal loans are consistent with a process of unilateral link formation. This, however, does not say anything about the nature of the willingness to link. In particular, do Nyakatoke villagers unilaterally decide *to whom* they wish to give? Or do they unilaterally decide *from whom* they can demand assistance?

The question on which proxies for willingness to link w_{ij} and w_{ji} are based is a question about an undirected link: *"Can you give a list of people [...] who you can personally rely on [...] and/or that can rely on you [...]?"* We do not know whether answers to this question capture willingness to provide help or to seek help – or both. But suppose we had separate information on i's willingness to give help to j and on i's willingness to ask j for help. Then we could test whether it is one or the other that drives the exchange of gifts and informal loans between Nyakatoke households.

To illustrate this idea, let w_{ij}^g denote i's desire to help j and let w_{ji}^r denote j's willingness to solicit help from i. Combining this information with the available information about the direction of transfers from i to j , we could construct a more specific test as follows: regress transfers τ_{ij} from i to j on w_{ij}^g and w_{ji}^r:

$$\tau_{ij} = \lambda(\alpha w_{ij}^g + \beta w_{ji}^r + \theta X_{ij})$$

If it is unilateral willingness to give that determines transfers, then we should have $\alpha > 0$ and $\beta = 0$: transfers take place whenever i wishes to give something to j. This could reflect altruism, or perhaps moral norms regarding charitable giving. In contrast, if it is unilateral willingness to receive help that determines τ_{ij}, transfers will take place whenever j wishes to receive something from i. Consequently we should obtain $\alpha = 0$ and $\beta > 0$. This could arise, for instance, because of social norms of redistribution, the existence of which has been argued by Hayami and Platteau (1996) for sub-Saharan Africa.[12]

[12]If j perfectly internalizes i's altruism towards him/her, then both α and β should in principle be positive. But since $w_{ji}^r = w_{ji}^g$ in this case, the $w_{ij}^g w_{ji}^r$ cross term will capture the effect of *both* w_{ij}^g and w_{ji}^r on transfers – and link formation will appear bilateral.

We do not have separate information about willingness to give and willingness to receive. But let us imagine for a moment that w_{ij} should in fact be interpreted as willingness to give, i.e., $w_{ij} = w^g_{ij}$. If this were the case, then when we regress τ_{ij} on w_{ij} and w_{ji}, it is like estimating a model of the form:

$$\tau_{ij} = \lambda(\alpha w^g_{ij} + \beta w^g_{ji} + \theta X_{ij})$$

If transfers are unilaterally driven by the willingness to give of the giver, then we should observe $\alpha > 0$ and $\beta = 0$. This is not what we observe in Tables 2 and 3.

Alternatively, imagine that answers to the undirected question of round 1 measure willingness to ask for help, i.e., $w_{ij} = w^r_{ij}$. In this case, when we regress τ_{ij} on w_{ij} and w_{ji}, it is like estimating a model of the form:

$$\tau_{ij} = \lambda(\alpha w^r_{ij} + \beta w^r_{ji} + \theta X_{ij})$$

If transfers are unilaterally driven by the recipient's willingness to request assistance, then we should observe $\alpha = 0$ and $\beta > 0$. Once again, this is not what we observe in Tables 2 and 3.

What inference can we draw from the above? First, there is no evidence that answers to the undirected question of round 1 should be interpreted as reflecting only willingness to give or only willingness to receive. If this had been the case, we should not have found w_{ij} and w_{ji} to be both significant in Tables 2 and 3 with coefficients of equal magnitude. It follows that answers to the undirected question of round 1 were indeed undirected: they capture both willingness to give and willingness to receive.

Secondly, we cannot a priori tell whether w_{ij} captures willingness to give and receive from the same person – as in a reciprocal relationship – or whether some w_{ij}'s capture willingness to give and others capture willingness to receive. But in the latter case, both types of w_{ij}'s would need to be present in the data in exactly the right proportions for α and β to be of equal magnitude. Since there is no particular reason for this to be the case, we find this possibility unlikely. It follows that w_{ij} most probably represents willingness to enter in a reciprocal relationship – as indeed suggested by the wording of the question.

6 Conclusion

Using detailed dyadic data from the village of Nyakatoke in Tanzania, we have tested whether gifts and loans between resident households is best interpreted as driven by unilateral or bilateral link formation. Two maintained assumptions underlie our analysis: answers to a first-round question on who people would turn to for help are good proxies for willingness to link; and reporting propensities are independent between giving and receiving households.

The first assumption is perhaps the most contentious one, given that the first-round question asks who respondents would turn to for help (or would provide help to), not who they would *like* to turn to for help. But there is considerable discordance between answers given by respondents, and detailed analysis by Comola and Fafchamps (2009) – and additional analysis conducted here – indicate that answers to this question are best interpreted as willingness to link. The second maintained assumption is also problematic, although it is weakened by the inclusion, in the two reporting equations, of regressors that capture some of the correlation in reporting. This assumption is only required for the estimation of the maximum likelihood model that explicitly recognizes the existence of mis-reporting – not for the estimation of a simpler, single-equation model.

The testing strategy itself is based on the simple observation that, if lending and gift giving are a voluntary agreement between two households, then both households should want to rely on each other for help. In contrast, if households can decide to form mutual assistance links with others they wish to help – or from whom they wish to seek assistance – then gifts and loans are best seen as a unilateral process.

We develop an estimation methodology that corrects for mis-reporting. This is essential for our purpose because there are substantial discrepancies between gifts and loans reported by givers and receivers. Simulation results indicate that, for the purpose of our test, ignoring reporting bias would lead to incorrect inference. We propose a new maximum likelihood estimator that corrects for response bias. Two versions of this estimator are developed: one in which we assume away over-reporting, and one in which we assume away under-reporting. In our data, it is unlikely that respondents systematically over-report the gifts and loans they give or receive. We therefore put more weight in the results that assume that respondents under-report gifts and loans – presumably because they forget.

When we assume that respondents under-report, we find no evidence to support bilateral link formation. We do, however, find reasonably convincing evidence to support the unilateral link formation hypothesis. These results are robust to different choices of model specification. When we assume that respondents over-report gifts and loans, the evidence is less conclusive: we still find evidence consistent with unilateral link formation, but some of the coefficient estimates are also consistent with bilateral link formation.

Given data limitations, we cannot formally test whether it is willingness to give or willingness to demand that drive transfers. But taken as a whole, the evidence is most consistent with transfers being driven by willingness to enter in a reciprocal relationship. If this interpretation is correct, the evidence of unilateral link formation that we have uncovered implies that if one household wishes to enter in a reciprocal relationship with another household, it can unilaterally do so – provided this other household is sufficiently close socially and geographically. This could arise, for instance, because inter-personal norms of reciprocity can be activated unilaterally by Nyakatoke villagers – as when giving to someone is a way of obligating him or her to reciprocate in the future (Platteau 2000).

If confirmed by future research, the above interpretation could explain the puzzling findings of Fafchamps and Gubert (2007) and those of De Weerdt and Fafchamps (2009) using the same data. These authors find that, contrary to theoretical predictions, households do not appear more likely to form risk sharing links with those who face less covariate risk. But if households can wait after shocks are realized before deciding who to ask for help, they need not worry about covariate risk *ex ante*.

This interpretation ties with another surprising result of our analysis, namely that loans are less likely to be reported than gifts. It is easy to see why borrowers would fail to report the loans they have received, but why would lenders do so? Much of the theoretical discourse about risk sharing has emphasized repeated games and reputation sanctions (Coate and Ravallion 1993, Kocherlakota 1996). Yet, if lenders hide the loans they make, it is hard to see how group reputational sanctions could be imposed. There must therefore be a cost to the lender for publicizing loans. One possible explanation is that lenders fear that disclosing loans reveals they have money they do not need, and would attract additional requests for help. A similar point is made by Anderson and Baland (2002) regarding secrecy within households to avoid claims on resources by spouses. If link formation was bilateral and assistance was voluntary, it would be possible to refuse to assist others and secrecy would not be necessary. These issues deserve further investigation.

References

[1] Anderson, Siwan and Jean-Marie Baland (2002), 'The Economics of Roscas and Intra-household Resource Allocation', *Quarterly Journal of Economics*, 117(3): 963-95

[2] Arcand, Jean-Louis and Marcel Fafchamps (2008), 'Matching in Community-Based Organizations', Department of Economics, Oxford (mimeograph)

[3] Christakis, Nicholas A. and James H. Fowler (2009), *Connected: The Surprising Power of Our Social Networks and How They Shape Our Lives*, Little, Brown and Company, London

[4] Coate, Stephen and Martin Ravallion (1993), 'Reciprocity Without Commitment: Characterization and Performance of Informal Insurance Arrangements', *Journal of Development Economics*, 40: 1-24

[5] Comola, Margherita and Marcel Fafchamps (2009), 'Testing Unilateral and Bilateral Link Formation', Department of Economics, Oxford (mimeograph)

[6] De Weerdt, Joachim and Stefan Dercon (2006), 'Risk-Sharing Networks and Insurance Against Illness', *Journal of Development Economics*, 81(2): 337-56, December 2006

[7] De Weerdt, Joachim and Marcel Fafchamps (2009), 'Social Networks and Insurance against Transitory and Persistent Health Shocks', Department of Economics, Oxford (mimeograph)

[8] Fafchamps, Marcel (1999), 'Risk Sharing and Quasi-Credit', *Journal of International Trade and Economic Development*, 8(3): 257-278

[9] Fafchamps, Marcel and Susan Lund (2003), 'Risk Sharing Networks in Rural Philippines', *Journal of Development Economics*, 71: 261-87

[10] Fafchamps, Marcel and Flore Gubert (2007), 'The Formation of Risk Sharing Networks', *Journal of Development Economics*, 83(2): 326-50

[11] Foster, Andrew D. and Mark R. Rosenzweig (2001), 'Imperfect Commitment, Altruism and the Family: Evidence from Transfer Behavior in Low-Income Rural Areas', *Review of Economics and Statistics*, 83(3): 389-407, August 2001

[12] Goyal, Sanjeev (2007), *Connections: An Introduction to the Economics of Networks*, Princeton University Press, Princeton and Oxford

[13] Jackson, Matthew O. (2009), *Social and Economic Networks*, Princeton University Press, Princeton

[14] Kocherlakota, Narayana R. (1996), 'Implications of Efficient Risk Sharing Without Commitment', *Rev. Econ. Stud.*, 63(4): 595-609, October 1996

[15] Ligon, Ethan, Jonathan P. Thomas, and Tim Worrall (2001), 'Informal Insurance Arrangements in Village Economies', *Review of Economic Studies*, 69(1): 209-44, January 2001

[16] Platteau, Jean-Philippe and Yujiro Hayami (1996), 'Resource Endowments and Agricultural Development: Africa vs. Asia', University of Namur and Aoyama Gakuin University, Tokyo. Paper presented at the IEA Round Table Conference The Institutional Foundation of Economic Development in East Asia, Tokyo, 16-19 December 1996.

[17] Platteau, Jean-Philippe (2000), *Institutions, Social Norms, and Economic Development*, Harwood Academic Publishers, Amsterdam

[18] Steglich, Christian E.G., Tom A.B. Snijders, and Michael Pearson (2010), 'Dynamic Networks and Behavior: Separating Selection from Influence', *Sociological Methodology*, (forthcoming)

[19] Udry, Christopher (1994), 'Risk and Insurance in a Rural Credit Market: An Empirical Investigation in Northern Nigeria', *Review of Economic Studies*, 61(3): 495-526, July 1994

Appendix

Table A1. Quintiles of declared loans and gifts

	Gifts		Loans	
Information given by:	giver	receiver	giver	receiver
nonzero obs.	996	824	350	237
cut-off values:				
0-20%	240	200	456	400
20-40%	500	450	900	700
40-60%	1000	850	1500	1532
60-80%	1796	1800	3000	3000
80-100%	39400	46800	60000	40000

Note: the total sample size is 14042 dyads, and the quintiles cut-off values are computed on nonzero observations only. Values expressed in *tzs*.

Table A2. Testing whether willingness to link

	dependent variable: w_{ij}
$popularity_i$	0.013
	(0.010)
$wealth_i$	0.009
	(0.007)
$popularity_j$	0.049***
	(0.004)
$wealth_j$	0.007**
	(0.003)
constant	-2.151***
	(0.061)

*** p<0.01, ** p<0.05, * p<0.1.
Estimator is Probit. Dyadic-robust standard errors in parentheses.

Table A3. Marginal effects

	$Pr(G = 1 \mid \tau = 1)$			
	gifts		loans	
	coeff.	mfx.	coeff.	mfx.
w^*_{ij}	1.4917	0.5388	0.5702	0.0425
$wealth_i$	-0.0354	-0.0123	-0.0411	-0.0018
$same\,religion^*$	0.0246	0.0086	-0.0580	-0.0025
$related^*$	0.4332	0.1634	-0.0795	-0.0032
$distance$	-0.5854	-0.2036	-0.0835	-0.0036
n_i	0.0259	0.0090	0.1126	0.0049
	$Pr(R = 1 \mid \tau = 1)$			
	gifts		loans	
	coeff.	mfx.	coeff.	mfx.
w^*_{ji}	1.9196	0.6625	1.2059	0.0707
$wealth_j$	-0.0447	-0.0125	-0.0120	-0.0002
$same\,religion^*$	0.0118	0.0033	-0.0409	-0.0006
$related^*$	0.6136	0.2092	0.1331	0.0022
$distance$	-0.5326	-0.1496	0.0201	0.0003
$female\ dependents_j$	-0.1495	-0.0420	-0.0467	-0.0006
$male\ dependents_j$	-0.1906	-0.0535	-0.2221	-0.0031

*dy/dx is for discrete change of dummy variable from 0 to 1

Table A4. Constant-only model

	gifts	loans
$\Pr(\tau = 1)$		
w_{ij}	3.222***	2.856***
	(0.396)	(1.074)
w_{ji}	3.749***	3.184**
	(0.534)	(1.457)
$w_{ij}w_{ji}$	13.490***	10.787***
	(1.037)	(2.972)
$wealth_i$	0.064***	0.036
	(0.013)	(0.029)
$wealth_j$	0.083**	0.021
	(0.036)	(0.017)
$same\,religion$	0.519***	0.208
	(0.120)	(0.155)
$blood\,link$	2.423***	0.950
	(0.359)	(0.635)
$distance$	-2.049***	-1.799***
	(0.466)	(0.432)
$hh\,members_i$	0.115***	0.065
	(0.040)	(0.109)
$hh\,members_j$	0.235*	0.106
	(0.123)	(0.094)
$constant$	-3.483***	-2.270***
	(0.388)	(0.610)
$\Pr(G = 1 \mid \tau = 1)$		
$constant$	0.143	-1.528***
	(0.225)	(0.403)
$\Pr(R = 1 \mid \tau = 1)$		
$constant$	-0.228*	-1.986***
	(0.133)	(0.320)

*** p<0.01, ** p<0.05, * p<0.1.
Dyadic-robust standard errors in parentheses.

Table A5. Results for total transfers

	(1)	(2)	(3)	(4)	(5)
	τ_{ij}^u	τ_{ij}^o	$\Pr(\tau = 1)$	$\Pr(G = 1\|\tau = 1)$	$Pr(R = 1\|\tau = 1)$
w_{ij}	1.392***	1.122***	2.414***	1.605***	
	(0.102)	(0.107)	(0.377)	(0.212)	
w_{ji}	1.565***	1.564***	2.793***		1.994***
	(0.093)	(0.100)	(0.311)		(0.202)
$w_{ij}w_{ji}$	-0.344*	-0.320*	0.510		
	(0.188)	(0.164)	(2.297)		
$wealth_i$	0.026***	0.011	0.076***	-0.039***	
	(0.007)	(0.007)	(0.016)	(0.012)	
$wealth_j$	0.031**	0.006	0.092*		-0.035***
	(0.015)	(0.010)	(0.052)		(0.012)
$same\,religion$	0.198***	0.170***	0.430*	0.058	0.078
	(0.052)	(0.062)	(0.256)	(0.212)	(0.182)
$related$	0.851***	0.558***	1.657**	0.527	0.708**
	(0.168)	(0.187)	(0.680)	(0.434)	(0.346)
$distance$	-0.839***	-0.937***	-1.593***	-0.667	-0.601
	(0.160)	(0.292)	(0.607)	(0.488)	(0.463)
$hh\,members_i$	0.045*	0.043*	0.100**		
	(0.024)	(0.023)	(0.049)		
$hh\,members_j$	0.109***	0.066**	0.272**		
	(0.040)	(0.032)	(0.113)		
n_i				0.053***	
				(0.016)	
$female\ dependents_j$					-0.139
					(0.125)
$male\ dependents_j$					-0.207*
					(0.123)
$constant$	-1.796***	-2.363***	-3.022***	-0.350	-0.357
	(0.143)	(0.162)	(0.505)	(0.533)	(0.333)

*** p<0.01, ** p<0.05, * p<0.1. Dyadic-robust standard errors in parentheses.