Universitat Pompeu Fabra, Barcelona

"Network Structure in Risk-Sharing Arrangements: Evidence from Rural Tanzania"

By Margherita Comola

January 2007

Abstract

Rural households in developing countries manage their exposure to risk and stabilize consumption through informal insurance arrangements. This paper approaches risk-sharing arrangements from a network perspective, investigating how risk-sharing networks are formed. When agents form a new link they also gain access to the larger network of partner's friends and friends of these friends. I test the hypothesis that agents not only take into account the wealth of their direct partners, but also consider the benefits they would get from the net of indirect contacts. A network formation model with full heterogeneity among agents is first presented, following Jackson and Wolinsky (1996), an estimation procedure is then proposed and applied to data on rural Tanzania. Results show that agents take into account not only potential partners' characteristics, but also their social connections. This suggests that risk sharing partners are not short-sighted, but actually consider the net advantage of potential links, evaluating also indirect benefits deriving from changes in their relative position with respect to all other agents. This paper contributes to both network theory and literature on risk sharing, in that it proposes an innovative procedure to estimate endogenous network formation models, and also provides evidence that indirect contacts have an explanatory value disregarded by all previous studies which are focused on direct relations only.

I would like to thank Antonio Cabrales, Marta Reynal and Rasa Karapandza. I also gratefully thank Joachim De Weerdt for sharing his data.

I. INTRODUCTION

As social capital theorists have stated long ago and economists have realized relatively recently, a person's net of contacts (family, friends and associates) constitutes an important disposable resource. This is indeed the case for developing countries. Whenever formal economic and financial institutions lack strength, individuals are forced to rely on informal relationships and interpersonal links assume an economic value. The most famous examples of this phenomenon are risk-sharing arrangements, such as rural community households using informal links to manage their exposure to risk and to stabilize consumption in the face of idiosyncratic shocks. This paper investigates how these risk-sharing arrangements are formed and whether the connection structure of the community affects the formation of links. Specifically, the question to be answered is: do agents choose their risk-sharing partners on the basis of these partners' characteristics only, or do they also attribute importance to the fact that new links give them access to the larger network of partner's friends, and friends of these friends? That is, are indirect contacts a relevant variable in the formation of risk-sharing arrangements?

A flourishing economic literature has studied informal insurance arrangements in rural communities both from a theoretical and from an empirical point of view. The majority of the empirical studies have shed light on the mechanisms behind the creation of links, to identify which variables help predict the existence of an insurance arrangement (Fafchamps and Lund (2003), De Weerdt (2004), Dekker (2004) and Udry and Conley (2004) among others). However, none of these studies have underlined the importance of the indirect contacts in the community network as a determinant of link formation. Doing so, they disregard the basic lesson from network theory, that is, the importance of the connection structure of the community. From the seminal research done by Jackson and Wolinsky (1996) and onwards, network theory, based on game theoretical reasoning, claims that not only direct contacts, but also the entire graph of indirect contacts, is relevant in the formation or severance of links. This paper is aimed to fill the hole between these two approaches: taking an endogenous network formation perspective, and using data on rural

Tanzania, I empirically investigate whether indirect connections have any importance in the process of link creation.

I first proceed from theory, setting up a framework consistent with the model by Jackson and Wolinsky (1996). In the model, agents form links among themselves, and links provide benefits and involve costs. All agents' decentralized decisions determine the structure of the network at the community level, and individual utility depends on the complete network structure. However, my framework differs from the majority of network situations in that it incorporates heterogeneity along both cost and benefit dimensions. Benefits from links increase in the wealth of potential partners and depend on the entire network structure. Costs of links are determined by the social distance between partners and by their income generating activities. From this benchmark model testable predictions to evaluate the importance of indirect contacts are derived, and the model is then estimated using data on the Tanzanian rural village of Nyakatoke.

Villages are the most common economic and social structure throughout the entire developing world, and by far the predominant reality in Sub-Saharan Africa. Two-thirds of Sub-Saharan Africans live in rural areas[1], and their economic and social life is determined within the borders of their village. Villages are typically composed by a small number of households whose income is primarily derived from agriculture; within the village there are no spatial barriers and information flows are smooth. Since most villages are located far form each other or in areas where transport is difficult, relations among individuals in the same village are frequent and complex, while relations with the exterior world are rare or inexistent. All these features, together with the lack of formal credit institutions and the massive presence of risk sharing arrangements, make villages the ideal setting to study endogenous network formation.

In rural communities all outcomes, from weddings to money lending, are determined by informal, multi-purpose interactions (Fafchamps and Lund (2003) and Hoddinott et al. (2005)). Risk-sharing activities also originate from socio-economic links. When households have no access to credit, either because no institutions provide it or

[1] Source: US EIA (on line at http://www.eia.doe.gov/emeu/cabs/subafricaenv.html).

because they cannot meet the collateral required to enter a formal transaction, private arrangements are used to stabilize consumption. These credits are normally used to finance primary needs like food or health-related expenses. In most cases a fixed repayment schedule was not specified at the time of receiving the loan and the interest rate is zero. As Fafchamps and Gubert (2005) point out, people do not form links specifically to pool income, but all informal loans originate from pre-existing interpersonal relationships[2]. Accordingly, in this paper I study mutual insurance agreements among agents, rather than loans and gifts that actually take place. In the Nyakatoke survey all adult individuals were 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 labour"*; this piece of information will be used in order to define whether a link exists and to trace the village architecture. I then propose and implement a procedure to test the effect of indirect contacts on link formation, and introduce new variables accounting for the gains in terms of indirect contacts. Results show that agents are not short-sighted but actually take into account the net advantage of potential links, evaluating also indirect benefits deriving from changes in their relative position with respect to all other agents. Data therefore suggests that network structure has its own importance, and that indirect contacts have an additional explanatory value disregarded by all previous studies, which are focused on direct relations only.

This paper contributes to both network theory and literature on risk sharing. First, it proposes an innovative procedure to estimate endogenous network formation models that can be applied to several other fields. The second major contribution is that it highlights the importance of indirect contacts as a determinant of risk sharing arrangements, which is crucial for the full understanding of the leading forces driving link formation.

None of the previous empirical studies on risk sharing explicitly recognize a role for network architecture itself, with the only exception of Krishnan and Sciubba (2005). However, while they theoretically derive the properties of equilibrium networks and

[2] The same perspective is adopted by Bramoullé and Kranton (2005) from a purely theoretical point of view: in their model individuals first set bilateral relations and then use these relations to share income after income shocks are realized; the network is given by the pattern of existing relations where agents commit to share income, rather than by the transfers themselves.

then test whether observed networks have these properties, I take a different approach, performing a structural analysis of the network formation model which, to the best of my knowledge, is entirely innovative. All previous studies on applied network formation take network structure and players' relative position as given, and assess their role as determinants of the social outcome (Calvò-Armengol, Patacchini and Zenou (2005), among others). I proceed in a different way, estimating the parameters of individual's utility function consistent with the network being pairwise stable. From the risk sharing point of view, the major lesson of this paper is that when agents choose the links they want to form they look not only at potential partners' income, but also at their "social" characteristics, that is, social connections (friends, and friends of these friends). Acknowledging the value of indirect contacts is a relevant step to the full understanding of the leading forces behind the creation of links. As Dasgupta (2003) points out, informal networks have effects that spill over to all areas of economic activity, and precisely for this reason it is crucial to understand the forces driving network formation. Understanding informal institutions is necessary to design policy interventions at the micro level. Without a good knowledge of the non-written rules driving informal ties, the design of social protection policies would simply result as inappropriate.

The paper is organized as follows. Section II contains a review on risk sharing literature. In Section III and IV the theoretical framework and the data are respectively presented. Section V explains the estimation procedure, while section VI presents empirical specifications and results. Section VI concludes summarizing the main findings. Finally, tables and figures are presented at the end of the paper.

II. LITERATURE REVIEW

As Mace (1991) has pointed out, when there are no private information and liquidity constraints the optimal insurance scheme is full income pooling. However, this theoretical benchmark does not apply to risk-sharing arrangements. In village economies, risk sharing is incomplete and insurance seems to take place not at the community level but among smaller groups (see Ravallion and Chauduri (1997) for India, Udry (1994) for Nigeria, and Fafchamps and Lund (2003) for Philippines).

Different explanations for this failure have been proposed: Ligon (1998) using data from rural India concludes that information asymmetry is the main obstacle to full risk sharing. Other authors such as Udry (1994) and Kocherlakota (1996) argue that, in rural villages, the assumption of full information is not unreasonable. Recent contributions apply the theory of limited commitment to justify incomplete insurance schemes observed in reality (Coate and Ravallion (1993) and Ligon, Thomas and Worrall (2000)). The network approach to risk sharing originates from the evidence that, in reality, insurance within villages is incomplete. Several empirical studies have tried to define the appropriate group for risk sharing and to shed light on the mechanisms through which these groups are created[3]. These studies explore the statistical features of observed links, and agree in pointing out kinship, friendship and neighbourhood as the main determinants of risk-sharing arrangements. Among them, Goldstein, De Janvry and Sadoulet (2002) use data on Ghana to trace the profile of people receiving credit and people excluded from credit. Fafchamps and Lund (2003) investigate how households deal with shocks in rural Philippines concluding that, due to imperfect commitment and information, mutual insurance doesn't take place at the village level, but in smaller groups of friends and relatives[4]. De Weerdt (2004) finds that the main variables predicting link formation are kinship, distance, religion and common friends; he also recognizes that income correlation is likely to generate inefficiencies and to affect negatively the outcome of risk sharing arrangements[5]. On the same line, Dekker (2004), studying network formation in rural Zimbabwe, identifies the types of social relation that are important to establish informal insurance ties. His contribution is mainly methodological in that he makes use of a dyadic model taking into account the dependence among observations[6]. A few contributions also explicitly focus on the impact of ethnicity: Grimard (1997) identifies ethnic ties as a possible risk-sharing group, and finds evidence of a partial consumption smoothing through ethnic lines, and Fafchamps (2003) investigates the role of ethnicity and networks in African domestic trade, finding no evidence of ethnic discriminations in

[3] Among the few theoretical contributions on risk-sharing networks, Bloch, Genicot and Ray (2004) characterize the properties of stable insurance schemes for exogenously given network structures, and Genicot and Ray (2003) study the effect of allowing subgroup deviation in risk sharing arrangements.
[4] Fafchamps and Lund (2003) recognize the importance of network shape in risk sharing arrangements. Since indirect links are poor bridges for money flows, they claim that under-connectivity in networks can be a source of imperfect and non-total risk sharing in the community.
[5] Also Fafchamps and Gubert (2005) recognize the detrimental role of this so-called "assortative matching".
[6] This is the so-called *p2* model proposed by Duijn, Snijders and Zijlstra (2004).

agricultural markets. Along similar lines, it is worth mentioning the descriptive assessments by Hoddinott, Dercon and Krishnan (2005), Rosenzweigh (1988), and Udry and Conley (2004).

None of these contributions explicitly recognize a role for network structure, with the only exception of Krishnan and Sciubba (2005) whose approach is by far the closest to the one I propose. They offer a bridge between the theoretical literature on endogenous network formation and the empirical work on informal networks, stressing the importance of both number of link *and* network architecture in determining the social outcome. Krishnan and Sciubba (2005) modify the co-author model by Jackson and Wolinsky (1996) allowing for heterogeneity among agents and provide a theoretical framework that yields testable predictions about the network architectures arising in equilibrium. . In their setting, farmers differ in productivity and decide with whom they want to form a link; *ceteris paribus*, a better-endowed farmer is a more appealing partner. All decentralized decisions determine the structure of equilibrium networks; since their model admits multiple equilibria, Krishnan and Sciubba identify the common features shared by any stable network architecture and then check whether labour sharing arrangements observed in rural Ethiopia are compatible with the model's prediction. My approach is analogous, with some important differences: I impose a less restrictive structure of externalities and allow for heterogeneity both with respect to costs *and* benefits. And, most importantly, Krishnan and Sciubba (2005) test whether the theoretical features of stable equilibria are consistent with empirical evidence, while I start from the observation of the equilibrium network to estimate the underlining parameters of the model.

In this paper I make an extensive use of the literature on endogenous network formation that has been flourishing in the last decade[7]. Models of network formation have originated primarily from two sources: the random graph literature by physicists (Guimera' et al. (2003) and Boguna' et al. (2004) among others) and the economic literature based on game theoretical reasoning. Some economists have approached network formation from a non-cooperative point of view (Bala and Goyal (2000), and Galeotti and Goyal (2002)). However, the majority of research papers focus on equilibrium networks, where links are formed at the discretion of self-interested

[7] See Jackson (2003) and Jackson (2005) for a review.

agents whose utility is given by the overall network structure. The analysis of equilibrium networks is based on both cooperative and non-cooperative considerations, and highlights the tensions between private incentives and overall efficiency. The reference model by Jackson and Wolinsky (1996) will be discussed in depth in Section III.

III. The THEORY

This section illustrates network games' basic notations, provides a review of the paper by Jackson and Wolinsky (1996) and presents the model to be estimated.

Let $N = \{1,...,n\}$ be a set of players connected in some network relationship. Links are the consequence of agreement between parts. A link is established and/or maintained only if there is joint consent. The network *g* describes which pairs of players are linked to each other. *g* is a list of unordered pairs of players $\{ij\}$, $i, j \in N$; $\{ij\} \in g$ indicates that players *i* and *j* are linked under the network *g*. For any network *g*, *g+ij* defines the network obtained by adding link *ij* to an existing network *g* and, analogously, *g-ij* defines the network obtained by deleting link *ij*. *N(g)* is the set of players with at least one link in the network *g*. Finally, the network g^N is the set of all subsets of *N* of size 2, usually defined as the "complete network", consequently, $G = \{g \subset g^N\}$ denotes the set of all possible networks on *N*.

A typical feature of network games is that the total utility generated, and the way it is allocated among players, depends on the network structure. The utility of each player not only depends on actions undertaken by his direct partners, but also on actions undertaken by all other agents. This is summarized by the value function and the allocation rule. Different network shapes generate different levels of utility, even if the set of players stays the same: the *value function* is a function $v : \{g | g \subset g^N\} \to R$ expressing the overall level of utility reached by the group of players for each network structure. The value function defines the efficiency benchmark. A network $g \subset g^N$ is (*strongly) efficient* if $v(g) \geq v(g')$ for all $g' \subset g^N$. On the other hand, the so-called allocation rule defines how this overall value is divided among players. That is, if we define $Y_i(g,v)$ as the payoff player *i* gets from graph *g* under the value function *V*, an *allocation rule* is a function $Y : G \times V \to R^N$ such that $\sum_i Y_i(g,v) = v(g)$ for all *g* and *v*. In this paper the allocation rule is simply the utility that players directly receive, accounting for both costs and benefits generated by the links they form, which is the natural allocation, ruling out posterior arrangements and side payments.

In order to identify which networks are likely to arise in various contexts, a notion of network stability has to be imposed. The "pairwise stability" by Jackson and Wolinsky (1996) is the first concept of network stability proposed in literature. Pairwise stability implies that the formation of a link requires the consent of both parties involved, while severance can be done unilaterally; formally, a network *g* is *pairwise stable* (*PWS*) if

(i) For all $ij \in g$, $Y_i(g,v) \geq Y_i(g-ij,v)$ and $Y_j(g,v) \geq Y_j(g-ij,v)$

(ii) For all $ij \notin g$, if $Y_i(g,v) < Y_i(g+ij,v)$ then $Y_j(g,v) > Y_j(g+ij,v)$

That is, a network is pairwise stable if, given the overall network structure, links which are profitable for both parties are actually formed, and each player does not benefit in severing any existent link. Pairwise stability does not depend on the process through which the network is formed. Moreover, it is a relatively weak concept since it only admits deviations on a single link at a time. Pairwise stability frequently admits large sets of stable allocations, which may result in the impossibility of drawing precise recommendations. Several refinements to restrict the set of stable equilibria have been proposed: for instance, group deviations (instead of pair deviations only) may be allowed as in Jackson and Van Den Nouweland (2005); alternatively, side payments between agents may also be implicitly allowed as in Jackson and Wolinsky (1996); however, for the purpose of this paper the pairwise stability concept will be adopted.

The "connection model" and the "cohautor model" by Jackson and Wolinsky (1996) are the first network situations presented in literature. In the connection model, agents decide whether to form links, which represent social relationships. Relationships provide benefits but also involve costs. Players incur a cost for every link they form;

on the other side, they benefit not only from direct (and therefore costly) relationships, but also from indirect ones, which are for free. Benefits from indirect relationships deteriorate with distance: a "friend" is more valuable than "a friend of a friend", which is more valuable than "a friend of a friend of a friend" and so on. Jackson and Wolinsky (1996) focus on a simplified, symmetric version of their general setting, assuming that every link has the same cost $c_{ij} = c$ and provides the same benefit normalized to 1. The payoff that player *i* receives from network *g* thus becomes

$$u_i(g) = \sum_{j \in N(g)} \delta^{t_{ij}} - \sum_{j:ij \in g} c_{ij}$$

with a depreciation rate $0 < \delta < 1$. t_{ij} is called the "*geodesic distance*" between *i* and *j* and is the number of links in the shortest path between *i* and *j* (setting $t_{ij} = \infty$ if there is no path between *i* and *j*). Thus, player *i* pays a cost *c* for direct connections only, but he also benefits from all indirect ones, in a way that is proportional to the proximity of these indirect partners. The value function is simply the sum of individual utilities $v(g) = \sum_{i \in N} u_i(g)$, and the allocation rule assigns to every player his own utility. As their main result, Jackson and Wolinsky (1996) illustrate the relationship between the sets of networks that are efficient and those that are stable, showing that these two sets do not generally coincide and decentralized decisions do not necessarily lead to an allocation that maximizes collective utility[8]. An analogous case is the one depicted by the "cohautor model", with the important difference that in this setting indirect connections provide negative (instead of positive) externalities, with a mechanism that will be explained in depth later on in Section VI.

In my setting I introduce full heterogeneity among agents, which results in differentiating benefits *and* costs. A similar setting has been analysed by Galeotti and Goyal (2002) from a non-cooperative perspective, however, up to my knowledge, pairwise stable networks with such general features have not been theoretically explored yet.

[8] In particular, for high and low costs efficient networks coincide with pairwise stable ones, while for the intermediate cases this may not be true.

In a community of $N = \{1,...,n\}$ agents each agent $i \in N$ is endowed with an income y_i, with a vector $\boldsymbol{z}_i$ of social characteristics (religion, ethnicity, blood links, schooling etc.), and with a vector $\mathbf{e}_i$ which describes his income source(s) out of a finite set of income generating activities. I assume that agents are more willing to form links with wealthy people and/or people that are closely connected to wealthy people: that is, the benefit from a link increases in the income of both direct and indirect potential partners; indirect connections are for free and become less valuable the more distant they are. The cost of linking is assumed to be increasing in the perceived social distance, which is based on innate or cultural characteristics such as religion, clan belonging, kinship and blood links: relationships between "similes" are relatively easier to form and maintain, and arrangements between partners with the same socio-cultural background have smaller enforcement and monitoring costs[9]. Finally, when agents choose whether to form a link, they also take into account the partners' income sources. Here, offsetting forces are at work: on one side two agents performing similar tasks tend to enter in contact and access private information, which facilitates the formation of a trust link. On the other side, in a fluctuating environment, insurance between agents with correlated income streams may be inefficient.

My objects of study are the unique pairs of agents called "dyads"; $l = \binom{n}{2}$ defines the dyads number. Let m, r and $s \in Z^+$ and using the information in $\boldsymbol{z}$ and $\boldsymbol{e}$ let us define a social attributes matrix $\boldsymbol{Z}_{(mxl)}$, where m are the attributes identifying the relative social position of each dyad, and a matrix $\boldsymbol{E}_{(rxl)}$ describing the productive portfolio and the income correlation for each dyad. Finally, let $\boldsymbol{Y}_{(sxn)}$ summarize each agent's wealth $\boldsymbol{y_i}$. Individual i's utility function from network g is therefore defined by

$$u_i(g) = y_i + \sum_{j \in N(g)} \delta(t_{ij})\boldsymbol{\alpha}\mathbf{y}_\mathbf{j} + \sum_{j:ij \in g} \boldsymbol{\beta}\mathbf{Z}_\mathbf{ij} + \sum_{j:ij \in g} \boldsymbol{\gamma}\mathbf{E}_\mathbf{ij}$$

Where:

$\delta : Z^+ \to R$ s.t. $\delta(.) \geq 0$, $\delta'(.) < 0$

t_{ij} is the geodesic distance between i and j

[9] Karlan (2001) applies a similar argument to microfinance group-lending programs in the Andes, and concludes that more homogeneous groups have higher repayment rates due to their higher social capital and ease of monitoring.

$\boldsymbol{\alpha} \in R^s$, $\boldsymbol{\beta} \in R^m$ and $\boldsymbol{\gamma} \in R^r$

In this formulation, agents' utility depends on the overall network structure, and externalities from indirect contacts can be positive or negative in sign. However, this is too general to draw precise conclusions: a "simple" equilibrium concept as pairwise stability is successful in pinning down specific network architectures under the assumption of homogeneous agents, but when agents face different incentives most predictive power is lost. I therefore proceed to empirics, and propose an estimation procedure to test the hypothesis that individual utility in risk sharing arrangements is affected by indirect contacts.

IV. The DATA

Data come from the Nyakatoke Household Survey. Nyakatoke is a small Haya[10] village in the Buboka Rural District of Tanzania, at the west of Lake Victoria. The community is composed by 600 inhabitants, 307 of which are adults, for a total of 119 households. Habitants have been interviewed in five regular intervals from February to December 2000. First all household heads, and a few days later, all adults were interviewed[11]; this has produced a rich dataset containing 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. Even if some information was collected at the individual level, the 7021 household dyads are taken as units of analysis.

Among other questions, adult households' members were 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 labour"*; this piece of information will be used in order to define whether a link exists and to trace the

[10] One of the largest tribes west of Lake Victoria.
[11] In order to eliminate possible sources of bias, gender sensitive issues were implemented by enumerators of the same sex as respondents.

village architecture. These links are reciprocal by definition, since people are asked to mention somebody they can rely on and/or that can rely on them. Informal links are also bilateral by their own nature, since they rely on agreements among parties involved and economic help is expected to be reciprocated at some point in the future. I thus assume links to be unweighted and undirected, and every time an individual mentions another one I draft a link between the two households they belong to[12]. With this procedure 490 links among the 119 households are identified[13]. The resulting network is dense, with a mean geodesic distance of 2.5 steps and a maximum geodesic distance of 5 steps. No household is isolated, and the number of households' reported links goes from 1 to 32. For a graphical representation of the network see Figure 1 in the Appendix. It also worth mentioning that Nyakatoke hosts more than 20 *formal* insurance groups, mostly aimed to help participants in the event of a funeral, which in Haya society is an important lump-sum expense. However, these groups follow a relatively rigid protocol in terms of acceptance, membership and contributions, therefore they cannot be compared with informal insurance arrangements and are out of the focus of analysis.

In order to define households' wealth I make use of two pieces of information: land and livestock[14] owned (both in Tanzanian shillings[15]). Relational variables between dyads are based on: religion, clan, geographical distance, schooling and kinship. All households in Nyakatoke nowadays follow modern religions being Muslim, Lutheran or Catholic[16]. Religion is a characterizing attribute in Nyakatoke also because the three main formal religious associations play an active role in the social and economic life of the village, providing help for funerals and giving support to the church or

[12] The entire exercise has also been repeated for the alternative definition of link: a link exists only if both households explicitly mention the partner. Results are consistent in sign and magnitude but not always in significance, which is not surprising given the exiguous number of observations (140 instead of 490).
[13] The nature of the exercise does not allow making use of links outside the village.
[14] Bulls, cows, calves, goats, sheep, pigs, chickens, and ducks.
[15] Data on land were originally in acres, but in order to allow comparisons they have been transformed in monetary equivalent with a conversion rate of 300000 tzs for 1 acre, which reflects average local prices in 2000.
[16] The first settlement in Nyakatoke dates around 1910: at that time most households were still adherents to traditional Bahaya religions. With time, however, the entire village was converted to modern religions. Presently the area north of the stream is predominantly Catholic; most of the Lutherans are in the south and Muslims southwest of the village. Since people converted after settlement, this distribution is likely to be a result of coincidence and mutual influence between neighbors and not willful segregation (Mitti and Rweyemamu (2001)).

mosque in cash, kind or labor (De Weerdt (2002)). Also clan belonging can potentially be a leading force of network formation: as De Weerdt (2004) acknowledges, *"(…) the clan is still an important institution in Haya culture, for example in matters regarding land rights. The clan elders can, in effect, function as a court of law. They could easily reprimand younger clan mates when they think their behavior is bad for the clan. Everybody wants to avoid falling out with their clan*". In Nyakatoke there are 26 different clans, with a variable number of households from 1 to 23. Also geographical distance is included in the empirical specification, in order to control for the fact that frequent interactions between neighbors can broaden their information and facilitate trust relations. However, the village area is relatively small, with an average distance between households of 523 meters and a maximum distance of 1738 meters. Information about schooling attainments is collected at individual level and classified in four categories: "no education", "started primary", "finished primary", and "secondary". In Nyakatoke there is no primary school (the closest one is at 2 km) and overall educational level is low, with 26 households out of 119 where no member has completed primary education. Education may impact link formation through several channels: on one side, it can be a dimension of similarity if households share the same educational attainments. On the other side, as De Weerdt (2004) suggests, education is a scarce and useful resource and households without literate members may find it interesting to befriend households with literate members. I finally take into account blood relations as a natural source of mutual insurance. I record a kinship tie between two households whenever a member of one household has a blood bond up to the third degree with a member of the other one, so that each dyad falls into one of the four categories: "parents, children and siblings", "Nephews, nieces, uncles, aunts, cousins, grandparents and grandchildren", "other blood bond", and "no blood bond".

Households in Nyakatoke get most of their income from agricultural activities, especially the cultivation of coffee and banana; other sources of income are rare and off-farming activities are mostly considered supplementary to farming[17]. During individual interviews, each active adult has listed the one or more productive activities he is engaged into. Productive activities are coded in seven categories: casual labour,

[17] Mitti and Rweyemamu (2001).

trade, crops, livestock rearing, assets, processing of agricultural products and other off-farm work.

Descriptive statistics are reported in the Appendix, Table 1-4. For additional information on Nyakatoke I remand to Mitti and Rweyemamu (2001) and De Weerdt (2002).

V. ESTIMATION PROCEDURE

For each pair of agents[18] *ij*, the dependent binary variable x_{ij} equals one if they are linked. Recalling the linear individual utility function

$$u_i(g) = y_i + \sum_{j \in N(g)} \delta(t_{ij}) \boldsymbol{\alpha} \mathbf{y}_\mathbf{j} + \sum_{j:ij \in g} \boldsymbol{\beta} \mathbf{Z}_\mathbf{ij} + \sum_{j:ij \in g} \boldsymbol{\gamma} \mathbf{E}_\mathbf{ij}$$

For an observed network *g*, α, β, γ`s have to be estimated under the constraint of pairwise stability, that is, imposing

$$\begin{aligned} \forall ij \in g, \quad & u_i(g) \geq u_i(g - ij) \\ & u_j(g) \geq u_j(g - ij) \\ \forall ij \notin g, \quad & u_i(g) < u_i(g + ij) \Rightarrow u_j(g) > u_j(g + ij) \end{aligned}$$

Under pairwise stability no subgroup deviations or multiple simultaneous deviations are allowed, and each agent considers whether to form/sever only one link at a time. Links are formed and maintained only if both agents involved agree. In order to decide whether the link *ij* is profitable each agents takes the equilibrium network *g* as given, and he *only* compares his utility under the two different scenarios when link *ij* is formed or not. That is, under pairwise stability agent *i* evaluates his utility from link *ij* taking the structure of *g* as *e*xogenous, all other links he is involved in (*ik* and *ki*,

[18] Whenever "agents" are mentioned it should be interpreted as "households".

$k \neq j$) and all other links player j is involved in (jk and kj, $k \neq i$) included. This *ceteris paribus* condition dramatically simplifies the estimation, since it rules out endogeneity. Therefore, for every dyad ij the model reduces to a discrete choice form

$$P(x_{ij}=1) = P\left(u_i(g) \geq u_i(g-ij) \;\&\; u_j(g) \geq u_j(g-ij)\right)$$

Where for each ij the regressors are calculated on the equilibrium network g when *only* the link ij varies. The utility of agent i under the two different scenarios is given by

$$z_{ij1} = y_i + \sum_{k \in N(g_{ij})} \delta(t_{ik})\boldsymbol{\alpha}\mathbf{y_k} + \sum_{k:ik \in g_{ij}} \boldsymbol{\beta}\mathbf{Z_{ik}} + \sum_{k:ik \in g_{ij}} \boldsymbol{\gamma}\mathbf{E_{ik}}$$

$$\bar{z}_{ij1} = y_i + \sum_{k \in N(g_{\bar{ij}})} \delta(t_{ik})\boldsymbol{\alpha}\mathbf{y_k} + \sum_{k:ik \in g_{\bar{ij}}} \boldsymbol{\beta}\mathbf{Z_{ik}} + \sum_{k:ik \in g_{\bar{ij}}} \boldsymbol{\gamma}\mathbf{E_{ik}}$$

And the utility of agent j by

$$z_{ij2} = y_j + \sum_{k \in N(g_{ij})} \delta(t_{jk})\boldsymbol{\alpha}\mathbf{y_k} + \sum_{k:jk \in g_{ij}} \boldsymbol{\beta}\mathbf{Z_{jk}} + \sum_{k:jk \in g_{ij}} \boldsymbol{\gamma}\mathbf{E_{jk}}$$

$$\bar{z}_{ij2} = y_j + \sum_{k \in N(g_{\bar{ij}})} \delta(t_{jk})\boldsymbol{\alpha}\mathbf{y_k} + \sum_{k:jk \in g_{\bar{ij}}} \boldsymbol{\beta}\mathbf{Z_{jk}} + \sum_{k:jk \in g_{\bar{ij}}} \boldsymbol{\gamma}\mathbf{E_{jk}}$$

where g_{ij} and $g_{\bar{ij}}$ are *ad hoc* network structures built for estimation purposes: to define g_{ij}, the architecture of g is taken as it is for all other dyads except ij, and additionally ij is assumed to exist. Analogously, $g_{\bar{ij}}$ is constructed by taking the rest of the network g as given and artificially setting $x_{ij}=0$ [19].

Adding the random disturbances,

[19] This is a slight abuse of notation, since g_{ij} should be defined as $g+ij$ if $x_{ij}=0$ and g if $x_{ij}=1$, and analogously $g_{\bar{ij}}$ should be defined as g-ij if $x_{ij}=1$ and g if $x_{ij}=0$. In everything that follows, g_{ij} and $g_{\bar{ij}}$ will refer to the artificial networks created from g setting a particular link ij to 1 or to 0 (regardless of the link being 1 or 0 in reality).

$$x_{ij}=1 \Leftrightarrow \begin{cases} z_{ij1}+\varepsilon_{ij1} \geq \bar{z}_{ij1}+\bar{\varepsilon}_{ij1} \\ z_{ij2}+\varepsilon_{ij2} \geq \bar{z}_{ij2}+\bar{\varepsilon}_{ij2} \end{cases}$$

And redefining

$$z_{ij1}-\bar{z}_{ij1}=\varphi_{ij1}$$

$$z_{ij2}-\bar{z}_{ij2}=\varphi_{ij2}$$

$$\varepsilon_{ij1}-\bar{\varepsilon}_{ij1}=\eta_{ij1}$$

$$\varepsilon_{ij2}-\bar{\varepsilon}_{ij2}=\eta_{ij2}$$

The model reduces to

$$P(x_{ij}=1)=P\,(\varphi_{ij1}+\eta_{ij1}\geq 0\ \ \&\,\varphi_{ij2}+\eta_{ij2}\geq 0)$$

Since η_{ij1} and η_{ji2} represent omitted factors affecting the utility that individuals *i* and *j* respectively get if link *ij* is formed, a joint distribution $\lfloor \eta_{ij1},\eta_{ij2} \rfloor$~*bivariate normal* $[0,0,1,1,\rho]$ is assumed[20]. The model is thus estimated as a *bivariate probit model with partial observability*. In the bivariate probit, two binary response variables vary jointly; partial observability occurs when a positive outcome is observed only if both response variables are positive. To proceed to the estimation, the problem is reformulated in the following way: let's define

$$\upsilon_{ij1}=\begin{cases}1 & if \quad \varphi_{ij1}+\eta_{ij1}\geq 0 \\ 0 & if \quad otherwise\end{cases} \quad \text{and} \quad \upsilon_{ij2}=\begin{cases}1 & if \quad \varphi_{ij2}+\eta_{ij2}\geq 0 \\ 0 & if \quad otherwise\end{cases}$$

Each of these dichotomous variables take value one when the corresponding agent, given the equilibrium network architecture, benefits from forming the link *ij*. What is observed is a binary variable x_{ij} that is the product of them:

[20] For instance I do not control for formal associations belongings and trade partnerships.

$$x_{ij} = \upsilon_{ij1} \cdot \upsilon_{ij2} = \begin{cases} 1 \;\; if \;\; \upsilon_{ij1}, \upsilon_{ij2} = 1 \\ 0 \;\; if \;\; otherwise \end{cases}$$

Now the problem reduces to

$$\begin{cases} P(\upsilon_{ij1} = 1) = P(\varphi_{ij1} + \eta_{ij1} \geq 0) \quad [eq1] \\ P(\upsilon_{ij2} = 1) = P(\varphi_{ij2} + \eta_{ij2} \geq 0) \quad [eq2] \end{cases}$$

Where errors are correlated, the observed outcome is $\upsilon_{ij1} \cdot \upsilon_{ij2}$ and linear restrictions are imposed such that the individual coefficients $\alpha, \beta, \gamma's$ in φ_{ij1} and φ_{ij2} are the same for Equation 1 and 2. This model is analogous to the one proposed by Poirier (1980) (see also Maddala (1983), Abowd and Farber (1982) and Farber (1983)).

For each dyad ij , benefits from linking are given by the additional utility that each agent respectively gets if the link *ij* is formed, compared to the case where the link is not formed. υ_{ij1}=1 if *ceteris paribus* player *i* finds it profitable to form link *ij*, that is, if

$$z_{ij1} - \bar{z}_{ij1} = y_i + \sum_{k \in N(g_{ij})} \delta(t_{ik}) \boldsymbol{\alpha} \mathbf{y_k} + \sum_{k:ik \in g_{ij}} \boldsymbol{\beta} \mathbf{Z_{ik}} + \sum_{k:ik \in g_{ij}} \boldsymbol{\gamma} \mathbf{E_{ik}} - \left[y_i + \sum_{k \in N(g_{\bar{ij}})} \delta(t_{ik}) \boldsymbol{\alpha} \mathbf{y_k} + \sum_{k:ik \in g_{\bar{ij}}} \boldsymbol{\beta} \mathbf{Z_{ik}} + \sum_{k:ik \in g_{\bar{ij}}} \boldsymbol{\gamma} \mathbf{E_{ik}} \right] =$$

$$= \left[\sum_{k \in N(g_{ij})} \delta(t_{ik}) \boldsymbol{\alpha} \mathbf{y_k} - \sum_{k \in N(g_{\bar{ij}})} \delta(t_{ik}) \boldsymbol{\alpha} \mathbf{y_k} \right] + \boldsymbol{\beta} \mathbf{Z}_{ij} + \boldsymbol{\gamma} \mathbf{E}_{ij} \geq 0$$

And analogously, υ_{ij2} =1 if

$$z_{ij2} - \bar{z}_{ij2} = y_j + \sum_{k \in N(g_{ij})} \delta(t_{jk}) \boldsymbol{\alpha} \mathbf{y_k} + \sum_{k:jk \in g_{ij}} \boldsymbol{\beta} \mathbf{Z_{jk}} + \sum_{k:jk \in g_{ij}} \boldsymbol{\gamma} \mathbf{E_{jk}} - \left[y_j + \sum_{k \in N(g_{\bar{ij}})} \delta(t_{jk}) \boldsymbol{\alpha} \mathbf{y_k} + \sum_{k:jk \in g_{\bar{ij}}} \boldsymbol{\beta} \mathbf{Z_{jk}} + \sum_{k:jk \in g_{\bar{ij}}} \boldsymbol{\gamma} \mathbf{E_{jk}} \right] =$$

$$= \left[\sum_{k \in N(g_{ij})} \delta(t_{jk}) \boldsymbol{\alpha} \mathbf{y_k} - \sum_{k \in N(g_{\bar{ij}})} \delta(t_{jk}) \boldsymbol{\alpha} \mathbf{y_k} \right] + \boldsymbol{\beta} \mathbf{Z}_{ij} + \boldsymbol{\gamma} \mathbf{E}_{ij} \geq 0$$

The term $\sum_{k \in N(g_{ij})} \delta(t_{ik}) \boldsymbol{\alpha} \mathbf{y_k} - \sum_{k \in N(g_{\bar{ij}})} \delta(t_{ik}) \boldsymbol{\alpha} \mathbf{y_k}$ represents the overall net gain in terms of discounted (direct and indirect) partners' income that agent *i* gets if link *ij* is formed with respect to the case where the link is not formed, and

$\sum_{k\in N(g_{ij})}\delta(t_{jk})\boldsymbol{\alpha}\mathbf{y_k} - \sum_{k\in N(g_{\overline{ij}})}\delta(t_{jk})\boldsymbol{\alpha}\mathbf{y_k}$ is the analogous for agent j.

According to my scope, I decompose these terms by geodesic distance:

$$\sum_{k\in N(g_{ij})}\delta(t_{ik})\boldsymbol{\alpha}\mathbf{y_k} - \sum_{k\in N(g_{\overline{ij}})}\delta(t_{ik})\boldsymbol{\alpha}\mathbf{y_k} = \boldsymbol{\alpha}\sum_{t_{ik}}\delta(t_{ik})[\underset{\substack{k\in N(g_{ij})\\ \delta(t_{ik})}}{\mathbf{y_k}} - \underset{\substack{k\in N(g_{\overline{ij}})\\ \delta(t_{ik})}}{\mathbf{y_k}}] =$$

$$\underbrace{\boldsymbol{\alpha}\cdot\delta(1)}_{\equiv\mathbf{a}_{i1}}\mathbf{y_j} + \underbrace{\boldsymbol{\alpha}\cdot\delta(2)}_{\equiv\mathbf{a}_{i2}}\underbrace{[\underset{\substack{k\in N(g_{ij})\\ \delta(t_{ik})=2}}{\mathbf{y_k}} - \underset{\substack{k\in N(g_{\overline{ij}})\\ \delta(t_{ik})=2}}{\mathbf{y_k}}]}_{\substack{2stepsGAIN\_i_{ij} \\ \left(\substack{\text{net gain of } i \text{ in terms}\\ \text{of income of agents 2 steps away}\\ \text{if link } ij \text{ is formed}}\right)}} + \underbrace{\boldsymbol{\alpha}\cdot\delta(3)}_{\equiv\mathbf{a}_{i3}}\underbrace{[\underset{\substack{k\in N(g_{ij})\\ \delta(t_{ik})=3}}{\mathbf{y_k}} - \underset{\substack{k\in N(g_{\overline{ij}})\\ \delta(t_{ik})=3}}{\mathbf{y_k}}]}_{\substack{3stepsGAIN\_i_{ij} \\ \left(\substack{\text{net gain of } i \text{ in terms}\\ \text{of income of agents 3 steps away}\\ \text{if link } ij \text{ is formed}}\right)}} + \ldots$$

and

$$\sum_{k\in N(g_{ij})}\delta(t_{jk})\boldsymbol{\alpha}\mathbf{y_k} - \sum_{k\in N(g_{\overline{ij}})}\delta(t_{jk})\boldsymbol{\alpha}\mathbf{y_k} = \boldsymbol{\alpha}\sum_{t_{jk}}\delta(t_{jk})[\underset{\substack{k\in N(g_{ij})\\ \delta(t_{jk})}}{\mathbf{y_k}} - \underset{\substack{k\in N(g_{\overline{ij}})\\ \delta(t_{jk})}}{\mathbf{y_k}}] =$$

$$\underbrace{\boldsymbol{\alpha}\cdot\delta(1)}_{\equiv\mathbf{a}_{j1}}\mathbf{y_j} + \underbrace{\boldsymbol{\alpha}\cdot\delta(2)}_{\equiv\mathbf{a}_{j2}}\underbrace{[\underset{\substack{k\in N(g_{ij})\\ \delta(t_{jk})=2}}{\mathbf{y_k}} - \underset{\substack{k\in N(g_{\overline{ij}})\\ \delta(t_{jk})=2}}{\mathbf{y_k}}]}_{\substack{2stepsGAIN\_j_{ij} \\ \left(\substack{\text{net gain of } j \text{ in terms}\\ \text{of income of agents 2 steps away}\\ \text{if link } ij \text{ is formed}}\right)}} + \underbrace{\boldsymbol{\alpha}\cdot\delta(3)}_{\equiv\mathbf{a}_{j3}}\underbrace{[\underset{\substack{k\in N(g_{ij})\\ \delta(t_{jk})=3}}{\mathbf{y_k}} - \underset{\substack{k\in N(g_{\overline{ij}})\\ \delta(t_{jk})=3}}{\mathbf{y_k}}]}_{\substack{3stepsGAIN\_j_{ij} \\ \left(\substack{\text{net gain of } j \text{ in terms}\\ \text{of income of agents 3 steps away}\\ \text{if link } ij \text{ is formed}}\right)}} + \ldots$$

in this way we can explicitly separate the net utility gains deriving from indirect contacts according to their distance. For instance, $[\underset{\substack{k\in N(g_{ij})\\ \delta(t_{ik})=2}}{\mathbf{y_k}} - \underset{\substack{k\in N(g_{\overline{ij}})\\ \delta(t_{ik})=2}}{\mathbf{y_k}}]$ expresses the net gain in terms of income of 2-steps-away contacts that agent i gets if the link ij is formed. This term can be positive or negative (for instance, it is negative when agent i sets a link with a new partner that used to be reachable in 2 steps, because the direct gain is reflected in a 2-steps-distance loss) and can be computed from the data available. This decomposition provides a straightforward way to test my hypothesis on data: if the a's coefficients turn out to be significant, this is an evidence that agents form risk sharing arrangement taking also into account the relative position and the income characteristics of indirect partners.

In our specification agent i's wealth is given by $\boldsymbol{y_i}$=($land_i$, $livestock_i$). Setting $\delta(t_{ij}) = 0$ for $t_{ij} \geq 4$ [21], we get

$$\sum_{k\in N(g_{ij})}\delta(t_{ik})\boldsymbol{\alpha}\mathbf{y_k} - \sum_{k\in N(g_{\bar{ij}})}\delta(t_{ik})\boldsymbol{\alpha}\mathbf{y_k} =$$

$$\mathbf{a}_{i11}land_{\mathbf{j}} + \mathbf{a}_{i12}\underbrace{[\underset{\substack{k\in N(g_{ij})\\ \delta(t_{ik})=2}}{land_{\mathbf{k}}} - \underset{\substack{k\in N(g_{\bar{ij}})\\ \delta(t_{ik})=2}}{land_{\mathbf{k}}}]}_{2stepsLAND\_GAIN\_i_{ij}} + \mathbf{a}_{i13}\underbrace{[\underset{\substack{k\in N(g_{ij})\\ \delta(t_{ik})=3}}{land_{\mathbf{k}}} - \underset{\substack{k\in N(g_{\bar{ij}})\\ \delta(t_{ik})=3}}{land_{\mathbf{k}}}]}_{3stepsLAND\_GAIN\_i_{ij}} +$$

$$\mathbf{a}_{i21}livestock_{\mathbf{j}} + \mathbf{a}_{i22}\underbrace{[\underset{\substack{k\in N(g_{ij})\\ \delta(t_{ik})=2}}{livestock_{\mathbf{k}}} - \underset{\substack{k\in N(g_{\bar{ij}})\\ \delta(t_{ik})=2}}{livestock_{\mathbf{k}}}]}_{2stepsLIVESTOCK\_GAIN\_i_{ij}} + \mathbf{a}_{i23}\underbrace{[\underset{\substack{k\in N(g_{ij})\\ \delta(t_{ik})=3}}{livestock_{\mathbf{k}}} - \underset{\substack{k\in N(g_{\bar{ij}})\\ \delta(t_{ik})=3}}{livestock_{\mathbf{k}}}]}_{3stepsLIVESTOCK\_GAIN\_i_{ij}}$$

And, analogously,

$$\sum_{k\in N(g_{ij})}\delta(t_{jk})\boldsymbol{\alpha}\mathbf{y_k} - \sum_{k\in N(g_{\bar{ij}})}\delta(t_{jk})\boldsymbol{\alpha}\mathbf{y_k} =$$

$$\mathbf{a}_{j11}land_{i} + \mathbf{a}_{j12}\underbrace{[\underset{\substack{k\in N(g_{ij})\\ \delta(t_{jk})=2}}{land_{\mathbf{k}}} - \underset{\substack{k\in N(g_{\bar{ij}})\\ \delta(t_{jk})=2}}{land_{\mathbf{k}}}]}_{2stepsLAND\_GAIN\_j_{ij}} + \mathbf{a}_{j13}\underbrace{[\underset{\substack{k\in N(g_{ij})\\ \delta(t_{jk})=3}}{land_{\mathbf{k}}} - \underset{\substack{k\in N(g_{\bar{ij}})\\ \delta(t_{jk})=3}}{land_{\mathbf{k}}}]}_{3stepsLAND\_GAIN\_j_{ij}} +$$

$$\mathbf{a}_{j21}livestock_{i} + \mathbf{a}_{j22}\underbrace{[\underset{\substack{k\in N(g_{ij})\\ \delta(t_{jk})=2}}{livestock_{\mathbf{k}}} - \underset{\substack{k\in N(g_{\bar{ij}})\\ \delta(t_{jk})=2}}{livestock_{\mathbf{k}}}]}_{2stepsLIVESTOCK\_GAIN\_j_{ij}} + \mathbf{a}_{j23}\underbrace{[\underset{\substack{k\in N(g_{ij})\\ \delta(t_{jk})=3}}{livestock_{\mathbf{k}}} - \underset{\substack{k\in N(g_{\bar{ij}})\\ \delta(t_{jk})=3}}{livestock_{\mathbf{k}}}]}_{3stepsLIVESTOCK\_GAIN\_j_{ij}}$$

The final specification I bring to the data is therefore a bivariate probit given by the two equations:

[21] Estimating the impact of contacts which are several steps away leads to technical difficulties: since the original network is densely connected, if we calculate the net gains for $t_{ij} \geq 4$ there are very few nonzero observations, and this leads to convergence problems in the maximum likelihood estimates. However, it is sensible to assume that agents do not benefit from partners that are really far from them, and however this does not affect the validity of the remaining results.

$$EQ1=\left[\begin{array}{l}\mathbf{a}_{i11}land_{j}+\mathbf{a}_{i12}\mathit{2stepsLAND\_GAIN\_i}_{ij}+\mathbf{a}_{i13}\mathit{3stepsLAND\_GAIN\_i}_{ij}+\\+\mathbf{a}_{i21}livestock_{j}+\mathbf{a}_{i22}\mathit{2stepsLIVESTOCK\_GAIN\_i}_{ij}+\mathbf{a}_{i23}\mathit{3stepsLIVESTOCK\_GAIN\_i}_{ij}+\\+\boldsymbol{\beta}_{i}\mathbf{Z}_{ij}+\boldsymbol{\gamma}_{i}\mathbf{E}_{ij}+\eta_{ij1}\geq 0\end{array}\right]$$

$$EQ2=\left[\begin{array}{l}\mathbf{a}_{j11}land_{i}+\mathbf{a}_{j12}\mathit{2stepsLAND\_GAIN\_j}_{ij}+\mathbf{a}_{j13}\mathit{3stepsLAND\_GAIN\_j}_{ij}+\\+\mathbf{a}_{j21}livestock_{i}+\mathbf{a}_{j22}\mathit{2stepsLIVESTOCK\_GAIN\_j}_{ij}+\mathbf{a}_{j23}\mathit{3stepsLIVESTOCK\_GAIN\_j}_{ij}+\\+\boldsymbol{\beta}_{j}\mathbf{Z}_{ij}+\boldsymbol{\gamma}_{j}\mathbf{E}_{ij}+\eta_{ij2}\geq 0\end{array}\right]$$

Where I constrain the coefficients to be the same for Equation 1 and 2:

$$\begin{array}{l}\mathbf{a}_{i11}=\mathbf{a}_{j11}\\\mathbf{a}_{i12}=\mathbf{a}_{j12}\\\mathbf{a}_{i13}=\mathbf{a}_{j13}\\\mathbf{a}_{i21}=\mathbf{a}_{j21}\\\mathbf{a}_{i22}=\mathbf{a}_{j22}\\\mathbf{a}_{i23}=\mathbf{a}_{j23}\\\boldsymbol{\beta}_{i}=\boldsymbol{\beta}_{j}\\\boldsymbol{\gamma}_{i}=\boldsymbol{\gamma}_{j}\end{array}$$

In order to compute the regressors *all* shortest paths among *all* players for *all* possible scenarios have to be calculated. That is, for each *ij*: the networks g_{ij} and $g_{\overline{ij}}$ are constructed, a matrix of geodesic distances is assigned to g_{ij} and $g_{\overline{ij}}$ respectively, and finally *all* $t_{ik}, t_{jk}\ \ \forall k : k \in N(g_{ij})$ and *all* $t_{ik}, t_{jk}\ \ \forall k : k \in N(g_{\overline{ij}})$ are calculated and multiplied for agents' land and livestock assets. This has been done with Dikstra's algorithm, a procedure to solve the single-source shortest path problem in graphs[22].

[22] This algorithm, proposed by Dijkstra (1959), was originally programmed for directed, weighted graphs. For a given pair of vertices of a graph, the algorithm finds the shortest path.

VI. SPECIFICATIONS and RESULTS

My units of observation are agents' dyads, and the binary dependent variable equals one if the two agents are connected by an undirected link. Every dyad is identified by two equations, one for each of the two households: I estimate a bivariate probit with partial unobservability and I constrain the coefficients to be the same for both equations. Five different specifications are described in what follows.

In the first specification (*Spec1*) the network structure is disregarded: agents take into account land and livestock assets for the specific partner under consideration only; additionally, a set of basic controls (geographical distance, kinship levels and clan belonging)[23] is included. In all other specifications, I include land and livestock gains from the entire network structure calculated as in Section V. The second specification (*Spec2*) includes land and livestock gains, and basic controls. The third specification (*Spec3*) includes land and livestock gains, basic controls and educational variables. A household is considered "educated" if at least one member has completed primary school, and "not educated" otherwise. Taking as reference the case of different educational levels, two dummies taking value one if both households are educated and both not educated respectively are included. The forth specification (*Spec4*) includes land and livestock gains, basic controls and religious belonging. Perceived distance between religious groups may not be always the same: since Catholics and Lutheran are both Christians[24], they are supposed to be ideologically closer between themselves than with Muslims. Therefore, dummies for every religious combination are included in order to capture each group's willingness to form links within its own religion and with people of other religions[25]. In the fifth specification (*Spec5*), educational and religious variables are jointly included. The sixth specification (*Spec6*) includes all previous variables, plus households' income generating activities. Here two mechanisms are potentially at work: on one side, income from a particular activity can be considered more valuable an asset. That is, it can be relevant also potential

[23] Geographical distance is expressed in meters, while same clan belonging is a dummy variable. Kinship is expressed by the dummies: "Strict kinship" (parents, children and siblings), "broad kinship" (nephews, nieces, uncles, aunts, cousins, grandparents and grandchildren) an "other blood bonds", where the default category is "no blood bonds".

[24] Barret (1982).

[25] Here the reference category is "both households are Lutherans".

partners' source of income, not only the income itself. Therefore, the total workforce share in the partner household devoted to each productive activity is included among the regressors [26]. On the other side, even though people engaged in similar activities form links in a natural way, they are also subject to similar income fluctuations, which make insurance arrangements less efficient. Even if the large majority of loans and gift are in the event of idiosyncratic shocks as sickness or death, villagers may realize the detrimental effect of income covariance in insurance arrangements, and therefore covariance terms for households' productive activities portfolios are also included[27].

Results are remarkably stable across all specifications. The most complete specification (*Spec6*) is presented and commented in what follows: Table 1 reports marginal effects, and Table 2 refreshes the variable's definition. For the entire set of results I remand to the Appendix, Table 5.

Direct partners' assets *land_j* and *livestock_j* are consistently positive and significant (even when these "traditional" variables are the only ones included, as in *Spec1*, Table V, in the Appendix): as expected, the richer a potential partner, the more desirable a link with him.

2steps_LIVESTOCK_GAIN and *3steps_LIVESTOCK_GAIN* represent the (positive or negative) net gains in term of livestock of 2 and 3-steps-away partners that the agent receives if he decides to form the link. These variables are constructed in accordance with network theory, and constitute an innovation in the empirical literature. In accordance with my intuition, they are significant and positive in sign. This evidence suggests that agents are not short-sighted but they actually take into account the net advantage of a new link, evaluating also indirect benefits deriving from changes in their relative position with respect to all other agents.

[26] This term is calculated as follows: in the survey individuals mention the one or more productive activities they are involved into, classifying them in seven categories (casual labour, trade, crops, livestock breeding, assets, processing of agri-products and other off-farm work). Since they do not mention the relative importance of each activity, individuals are assumed to equally divide their time among all activities they are involved in. A household's total workforce corresponds to the number of its active members, and the shares of total workforce devoted to each activity are calculated. In case an individual does not earn any income, the information is coded as a zero (note that 5 households have no member earning any income).

[27] For each dyad and for each productive activity, the product the two households' shares is used as measure of overlapping by sector.

Table 1: Bivariate Probit Marginal Effects[28]

VARIABLE	dy/dx	P>|z|
land_j	.0020332	0.003
livestock_j	.0125078	0.000
2steps_LAND_GAIN	-.0012536	0.000
3steps_LAND_GAIN	-.0004188	0.058
2steps_LIVESTOCK_GAIN	.0062917	0.002
3steps_LIVESTOCK_GAIN	.0057006	0.005
distance (mts)	-.0000631	0.000
strict_kinship	.2131954	0.002
broad_kinship	.0798821	0.001
Other blood links	.0480781	0.001
Same clan	.0002623	0.957
Religion: both Muslim	.0219675	0.019
Religion: both Catholic	-.0085584	0.071
Religion: Lutheran + Catholic	-.0057569	0.179
Religion: Catholic+Muslim	-.0228431	0.000
Religion: Lutheran + Muslim	-.0069864	0.140
Education: both low	-.0033343	0.648
Education: both high	.0041251	0.203
Workforce share in off-farm	.0378192	0.071
Workforce share in casual labour	.0280453	0.122
Workforce share in trade	.0010232	0.957
Workforce share in cropping	.0271912	0.260
Workforce share in livestock	.0181362	0.427
Workforce share in assets	.1889668	0.011
Workforce share in processing	.0635219	0.003
Covariance off-farm	.0317597	0.504
Covariance casual labour	-.0499187	0.083
Covariance trade	.0001152	0.998
Covariance cropping	-.0217649	0.388
Covariance livestock	-.1220129	0.229
Covariance assets	1.925.709	0.197
Covariance processing	.0065028	0.882

[28] For *land_j*, *livestock_j* and workforce shares in each productive activity: means are calculated on the original sample of 119 households. For *2steps_LAND_GAIN, 3steps_LAND_GAIN, 2steps_LIVESTOCK_GAIN, and 3steps_LIVESTOCK_GAIN*: means are calculated on all households in the dyads (7021+7021 observations). All other variables' means are calculated on the sample of 7021 dyads.

Table 2: Variables Definitions

land_j	*Monetary value of land owned by household j (1 unit=100000 tzs)*
livestock_j	*Monetary value of livestock owned by household j (1 unit=100000 tzs)*
2steps_LAND_GAIN	*Net gain in term of land of 2 steps away agents if link ij is formed (1 unit=100000 tzs)*
3steps_LAND_GAIN	*Net gain in term of land of 3 steps away agents if link ij is formed (1 unit=100000 tzs)*
2steps_LIVESTOCK_GAIN	*Net gain in term of livestock of 2 steps away agents if link ij is formed (1 unit=100000 tzs)*
3steps_LIVESTOCK_GAIN	*Net gain in term of livestock of 3 steps away agents if link ij is formed (1 unit=100000 tzs)*
distance (mts)	*Distance between the households' houses (mts)*
strict_kinship	*Dummy variable, equals 1 if adults in the two households are respectively parents, children or siblings*
broad_kinship	*Dummy variable, equals 1 if adults in the two households are respectively nephews, nieces, uncles, aunts, cousins, grandparents or grandchildren*
Other blood links	*Dummy variable, equals 1 for blood bonds other than the ones above*
Same clan	*Dummy variable, equals 1 if the two households belong to the same clan*
Religion: both Muslim	*Dummy variable, equals one if both households are Muslim*
Religion: both Catholic	*Dummy variable, equals one if both households are Catholic*
Religion: Lutheran + Catholic	*Dummy variable, equals one if households are Catholic and Lutheran respectively*
Religion: Catholic+Muslim	*Dummy variable, equals one if households are Catholic and Muslim respectively*
Religion: Lutheran + Muslim	*Dummy variable, equals one if households are Lutheran and Muslim respectively*
Education: both low	*Dummy variable, equals one if in neither of the two households a member has completed primary school*
Education: both high	*Dummy variable, equals one if in both households at least one member has completed primary school*
Workforce share in off-farm	*Share of total workforce devoted to off-farm labour in household j*
Workforce share in casual labour	*Share of total workforce devoted to casual labour in household j*
Workforce share in trade	*Share of total workforce devoted to trade in household j*
Workforce share in cropping	*Share of total workforce devoted to cropping in household j*
Workforce share in livestock	*Share of total workforce devoted to livestock rearing in household j*
Workforce share in assets	*Share of total workforce devoted to assets in household j*
Workforce share in processing	*Share of total workforce devoted to processing in household j*
Covariance off-farm	*(workforce share in off-farm)i*(workforce share in off-farm)j*
Covariance casual labour	*(workforce share in casual labour)i*(workforce share in casual labour)j*
Covariance trade	*(workforce share in trade)i*(workforce share in trade)j*
Covariance cropping	*(workforce share in cropping)i*(workforce share in cropping)j*
Covariance livestock	*(workforce share in livestock)i*(workforce share in livestock)j*
Covariance assets	*(workforce share in assets)i*(workforce share in assets)j*
Covariance processing	*(workforce share in processing)i*(workforce share in processing)j*

On the other side the analogous variables for land assets (*2steps_LAND_GAIN* and *3steps_LAND_GAIN*) are significant across most specifications, but their sign is unexpectedly negative. Data seem thus to suggest that land assets for direct contacts are favourably taken into account, while for indirect contacts the opposite holds. This implies that land and livestock are perceived differently: indirect livestock assets have a positive value; and in contrast for land assets the value is negative. Various interpretations of the phenomenon can be proposed, but no conclusive explanation can be provided; however it is worth noticing that the feature is remarkably stable across all specification.

In the Nyakatoke sample, the distribution of land is relatively egalitarian (see Figure 2 in the Appendix). According to Mitti and Rweyemamu (2001) “Almost every family in Nyakatoke owns the land it lives on and cultivates”, which is mainly devoted to satisfy households primary needs. On the other side, livestock is more unequally distributed (see Figure 3, Appendix) and thus it represents a significant wealth dimension. Being for its unequal distribution or for the fact that it is more disposable of an asset and therefore particularly useful in small contingent loans, livestock is worth dramatically more than land for risk sharing purposes (see Table 1: for the same units of measure, coefficients for livestock are consistently greater than the ones for land assets). While livestock assets of indirect partners are positively perceived, for land assets we can imagine that a “competition mechanism” analogous to the one driving the “cohautor model” by Jackson and Wolinsky (1996) prevails on the beneficial effects of indirect connections. Assume that each household has a limited amount of agricultural resources to devote to loans and gifts, because land is not a liquid asset and the household itself needs most of its production. Since everyone prefers to invest his limited resources in helping wealthier partners, benefits from a link are inversely correlated to the number and the assets of the partner’s friends, and therefore externalities from indirect contacts may result negative.

All other variables have the expected sign, and reconfirm the evidence documented by previous studies on risk sharing networks (Fafchamps and Lund (2003), De Weerdt (2004), Dekker (2004)). Geographical proximity is significantly correlated with the existence of a link between two households. The same holds for kinship ties; the stricter the kinship tie the stronger the effect. The likelihood of observing a link is greater when households' components are respectively parents and children than in the case were they are grandparents and grandchildren. On the other side, clan belonging and educational attainment do not seem to be relevant for link formation. Looking at religious belonging, we notice that Muslim households are much more willing to form links with people of their same religion than Catholic and Lutherans respectively are. That is, Muslims have a much stronger positive bias towards other Muslims, which can be interpreted in the light of their ideological distance from the two Christian groups. Also this can result from a "minority effect" since Muslims are the smallest of the three religious groups. Additionally, the likelihood of observing a link between Catholics and Muslims is significantly smaller than for any other religious combination. Regarding income sources, data suggest that only assets and processing have significant coefficients, that is, are considered more valuable for risk-sharing purposes, while all variables accounting for activity overlapping seem not to impact the process of link formation.

VII. CONCLUSIONS

This paper approaches risk-sharing arrangements in rural villages from a network perspective, investigating how risk-sharing networks are formed and whether indirect contacts are relevant in the process of link formation. When agents form a link not only do they establish a new contact, but they also gain access to the larger network of the partner's friends and friends of these friends. In this paper I test the hypothesis that agents not only take into account the wealth of direct partners, but they also consider the benefits they would get from the net of indirect contacts if they actually form a new link. A network formation model with full heterogeneity among agents is first presented following Jackson and Wolinsky (1996); an estimation protocol is then proposed and applied to data on rural Tanzania. Results show that agents are not short-sighted, but they actually take into account the net advantage of potential links, evaluating also indirect benefits deriving from changes in their relative position with respect to all other agents. This paper contributes to both network theory and literature on risk sharing, in that it proposes an innovative procedure to estimate endogenous network formation models, and also provides evidence that indirect contacts have an explanatory value disregarded by all previous studies, which are focused on direct relations only.

REFERENCES

Abowd, J.M and H.S. Farber (1982): "Job Queues and the Union Status of Workers" *Industrial and Labour Relations Review*, 35, 3, 354-367.

Bala, V. and S. Goyal (2000): "A Noncooperative Model of Network Formation", *Econometrica*, 68, 5, 181-1229.

Barret, D. (1982): "*World Christian Encyclopaedia*", Oxford University Press.

Bloch, F., G. Genicot and D. Ray (2004): "Informal Insurance in Social Networks" Georgetown University Department of Economics Working Paper 04-04-16.

Bogugna', M., R. Pastor-Satorras , A. Diaz-Guilera and A. Arenas (2004): "Models of Social Networks based on Social Distance Attachment", *Physical Review* 70.

Bramoulle', Y. and R. Kranton (2005): "Risk Sharing Networks", CIRPEE Working Paper 526.

Calvó-Armengol, A., E. Patacchini and Y. Zenou (2005): "Peer Effects and Social Networks in Education and Crime", Mimeo.

Coate, S. and Ravallion, M. (1993): "Reciprocity without Commitment: Characterization and Performance of Informal Insurance Arrangements", *Journal of Development Economics*, 40, 1-24.

Dasgupta, P. (2003): "Social Capital and Economic Performance: Analytics" published in *Foundations of Social Capital*, E. Ostrom and T.K. Ahn, eds (Cheltenham UK: Edward Elgar).

Dekker, M. (2004): "Risk sharing in rural Zimbabwe: An empirical Analysis of Endogenous Network Formation", Working Paper.

De Weerdt, J. (2002): “Community Organisations in Rural Tanzania: A Case Study of the Community of Nyakatoke, Buboka Rural District” Mimeo.

De Weerdt, J. (2004): “Risk Sharing and Endogenous Network Formation”, *Insurance Against Poverty*, Oxford University Press, chapter 10.

Dijkstra, E.D (1959): “A note on two problems in connexion with graphs”, *Numerische Mathematik*. 1, 269–271.

Duijn M., T. Snijders and B. Zijlstra (2004): “p2: A Random Effect Model with Covariates for Directed Graphs” , *Statistica Neerlandica*, 58, 2, 234-254(21).

Fafchamps, M. (2003): "Ethnicity and Networks in African Trade", *Contributions to Economic Analysis & Policy*: Vol. 2: No. 1, Article 14.

Fafchamps, M. and F. Gubert (2005): “The Formation of Risk Sharing Networks” DIAL Working Paper DT2005-13.

Fafchamps, M. and S. Lund (2003): “Risk Sharing Networks in Rural Philippines” *Journal of Development Economics*, 71,2, 261-287 (27).

Farber, H.S. (1983): “The Determination of Union Status of Workers”, *Econometrica*, 51 (5), 1417-1438.

Galeotti, A. and S. Goyal (2002): “Network Formation with Heterogeneous Players”, Tinbergen Institute Discussion Paper No. 2002-069/1.

Genicot, G. and D. Ray (2003): “Group Formation in Risk Sharing Arrangements”, *Review of Economic Studies* 70, 87-113.

Goldstein, M., De Janvry, A. and E. Sadoulet (2002) “Is a Friend in Need a Friend Indeed? Inclusion and Exclusion in Mutual Insurance Networks in Southern Ghana”, STICERD Working Paper DEDPS 32.

Grimard, F. (1997): “Households Consumption Smoothing through Ethnic Ties: Evidence form Core D’Ivoire” *Journal of Development Economics*, 53, 391-422.

Guimera’, R., L. Danon, A. Diaz-Guilera, F. Giralt, and A. Arenas (2003): “Self-similar Community Structure in a Network of Human Interactions”, *Physical Review* 68.

Hoddinott, J., S. Dercon and P. Krishnan (2005): “Networks and Informal Mutual Support in 15 Ethiopian Villages”, Mimeo.

Jackson, M. (2003): “A Survey of Models of Network Formation: Stability and Efficiency”, CALTECH Humanities and Social Sciences Division Working Paper 1161.

Jackson, M. (2005): “The Economics of Social Networks”, CALTECH Humanities and Social Sciences Division Working Paper 1237.

Jackson, M. and A. Van Den Nouweland (2005): “Strongly Stable Networks”, *Games and Economic Behaviour*, 51 (2), 420-444.

Jackson, M and A. Wolinsky (1996): “A Strategic Model of Social and Economic Networks”, *Journal of Economic Theory*, 71,44-74.

Karlan, D. S. (2001): “Social Capital and Group Banking”, CMFNET Working Paper.

Kocherlakota, N.R. (1996): “Implications of Efficient Risk Sharing without Commitment”, *Review of Economic Studies*, 65, 847-864.

Krishnan, P. and E. Sciubba (2005): “Links and Architecture in Village Networks”, Cambridge Working Papers in Economics 0462.

Ligon, E. (1998): “Risk Sharing and Information in Village Economies”, *Review of Economic Studies* 65, 847-864.

Ligon, E., J.P Thomas and T. Worrall (2000): "Mutual Insurance, Individual Savings and Limited Commitment", *Review of Economic Dynamics*, 3, 216-246.

Mace, B.J. (1991): "Full Insurance in the Presence of Aggregate Uncertainty", *Journal of Political Economy*, 99(5), 928-956.

Maddala, G.S. (1983): "*Limited Dependent and Qualitative Variables in Econometrics*", Cambridge University Press.

Mitti R. and T. Rweyemamu (2001): "A Socio-Economic Profile of Nyakatoke", Mimeo.

Poirer, D.J. (1980): "Partial Observability in Bivariate Probit Models" *Journal of Econometrics*, 12, 209–217.

Ravallion, M. and Chaudhuri, S. (1997): "Risk and Insurance in Village India: Comment", *Econometrica*, 65(1), 171-184.

Rosenzweigh, M.R. (1988): "Risk, Implicit Contract and the Family in Rural Areas of Low-income Countries" *The Economic Journal*, 98, 393, 1148-1170.

Udry, C. (1994): "Risk and Insurance in a Rural Credit Market: an Empirical Investigation in Northern Nigeria", *Review of Economic Studies*, 61, 495-526.

Udry, C.R. and T.G. Conley (2004): "Social Networks in Ghana" Yale University Economic Growth Centre Discussion Paper n. 888.

A. APPENDIX

Table 1: Nyakatoke population structure

Age category	***Male***	***Female***	***Total***
younger than 10	93	104	197
10 to 20	75	74	149
20 to 30	38	50	88
30 to 40	29	29	58
40 to 50	22	23	45
50 to 60	14	14	28
60 to 70	9	7	16
70 to 80	3	8	11
older than 80	3	5	8
TOTAL	286	314	600

Source: De Weerdt (2004)

Table 2: Distribution of dyadic variables

variable	***definition***	***distribution across the 7021 dyads***
Link	No link	6531
	Unilateral or reciprocal link	490
Geodesic distance (shorthest distance between the households on the network graph)	1 step	490
	2 steps	1996
	3 steps	2900
	4 steps	1275
	5 steps	360
Kinship	Child, siblings, parents	109
	Nephew, niece, auncle, aunt, cousins, grandparents, grandchildren	102
	Other blood links	172
	No blood link	6638
Education	Both households have at least one memeber who completed primary education	4278
	Only one household has at least one member who completed primary education	2418
	None of the two households have any member who completed primary education	325

Source: Nyakatoke Household Survey

Table 3: Distribution of attribute variables among households

Religion	Muslim Lutheran Catholic	24 46 49	n=119
Clan	1 household 2 households 3 households 4 to 10 households 12 to 23 households	11 5 2 5 3	n=119
Income generating activities: number of households engaged in each activity	casual labour trade crops livestock assets processing other off-farm	57 41 108 31 8 45 40	n=116

Source: De Weerdt (2004) and Nyakatoke Household Survey

Table 4: Land and livestock households' assets by quintile

Quintile	*Mean livestock value (Tsh)*	*Mean land value (Acres)*
1st	0	0.29
2nd	1867	0.62
3rd	7400	0.98
4th	23544	1.45
5th	254018	3.45
TOTAL	53354	1.34

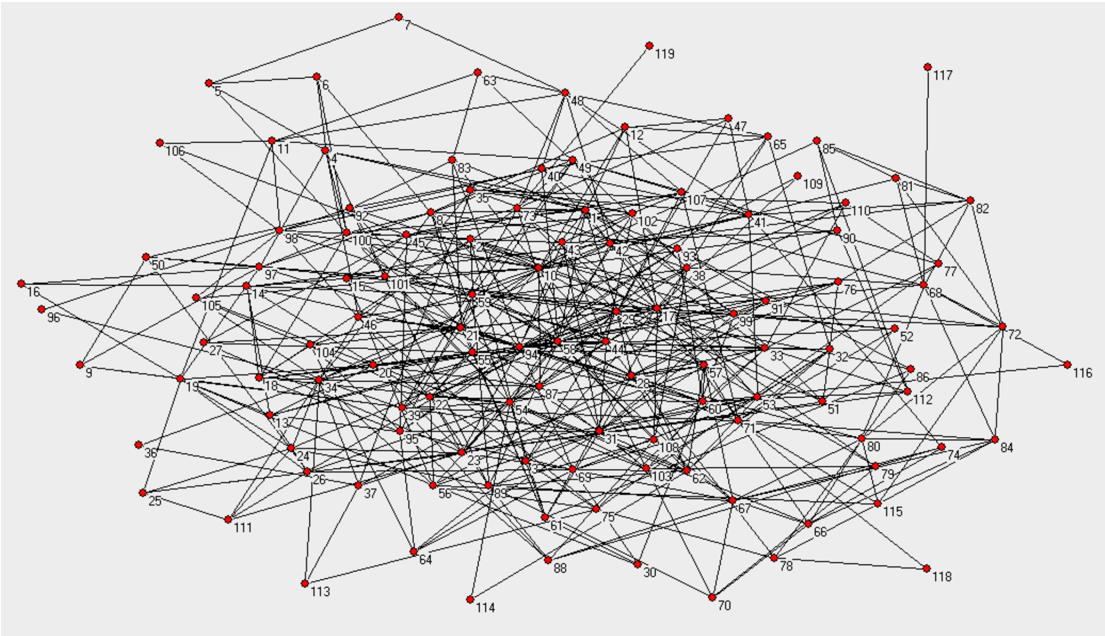
Note: original data on land in acres are used.
Source: De Weerdt (2004) and Nyakatoke Household Survey

Table 5: Results

VARIABLE	Spec1	Spec2	Spec3	Spec4	Spec5	Spec6
land_j	0.065011	0.025004	0.023016	0.030269	0.027367	0.017145
	(0.000)**	(0.008)**	(0.009)**	(0.012)*	(0.011)*	(0.012)*
livestock_j	0.127336	0.1036	0.1065	0.1278	0.1302	0.1055
	(0.033)*	(0.007)**	(0.005)**	(0.006)**	(0.003)**	(0.005)**
2steps_LAND_GAIN		-0.012067	-0.012369	-0.014129	-0.014336	-0.010571
		(0.003)**	(0.003)**	(0.002)**	(0.001)**	(0.004)**
3steps_LAND_GAIN		-0.004962	-0.004935	-0.006385	-0.006220	-0.003532
		(0.047)*	(0.047)*	(0.030)*	(0.030)*	(0.097)
2steps_LIVESTOCK_GAIN		0.0609	0.0628	0.0749	0.0763	0.0531
		(0.017)*	(0.014)*	(0.010)**	(0.008)**	(0.020)*
3steps_LIVESTOCK_GAIN		0.0550	0.0563	0.0687	0.0693	0.0481
		(0.023)*	(0.020)*	(0.014)*	(0.011)*	(0.027)*
distance (mts)	-0.001124	-0.001	-0.001	-0.001	-0.001	-0.001
	(0.000)**	(0.002)**	(0.001)**	(0.002)**	(0.001)**	(0.001)**
strict_kinship	16.496	1.072	1.078	1.164	1.162	0.987
	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**	(0.000)**
broad_kinship	0.9191	0.526	0.533	0.609	0.613	0.466
	(0.000)**	(0.001)**	(0.001)**	(0.001)**	(0.001)**	(0.001)**
Other blood links	0.6504	0.375	0.380	0.387	0.387	0.313
	(0.000)**	(0.001)**	(0.001)**	(0.002)**	(0.001)**	(0.001)**
Same clan	0.088	0.038	0.037	0.027	0.025	0.002
	(0.257)	(0.421)	(0.436)	(0.608)	(0.635)	(0.957)
Religion: both Muslim				0.194	0.187	0.162
				(0.040)*	(0.043)*	(0.031)*
Religion: both Catholic				-0.027	-0.037	-0.076
				(0.620)	(0.511)	(0.114)
Religion: Lutheran + Catholic				-0.027	-0.032	-0.049
				(0.576)	(0.510)	(0.207)
Religion: Catholic+Muslim				-0.280	-0.290	-0.221
				(0.010)**	(0.006)**	(0.005)**
Religion: Lutheran + Muslim				-0.099	-0.104	-0.061
				(0.123)	(0.106)	(0.188)
Education: both low			-0.038		-0.052	-0.029
			(0.609)		(0.554)	(0.660)
Education: both high			0.040		0.050	0.035
			(0.219)		(0.186)	(0.229)
Workforce share in off-farm						0.319
						(0.086)
Workforce share in casual labour						0.236
						(0.123)
Workforce share in trade						0.009
						(0.957)
Workforce share in cropping						0.229
						(0.272)
Workforce share in livestock						0.153
						(0.429)
Workforce share in assets						1.593
						(0.034)*
Workforce share in processing						0.536
						(0.009)**
Covariance off-farm						0.268
						(0.507)
Covariance casual labour						-0.421
						(0.100)
Covariance trade						0.001
						(0.998)
Covariance cropping						-0.184
						(0.402)
Covariance livestock						-1.029
						(0.232)
Covariance assets						16.239
						(0.199)
Covariance processing						0.055
						(0.881)
Constant	-0.781	0.048	0.073	0.031	0.060	-0.332
	(0.000)**	(0.896)	(0.848)	(0.933)	(0.873)	(0.379)
Observations	6670	6670	6670	6670	6670	6670

p values in parentheses
* significant at 5%; ** significant at 1%

Figure 1: Links among Nyakatoke households



Drawn using Pajek.

Figure 2: Land Quantiles

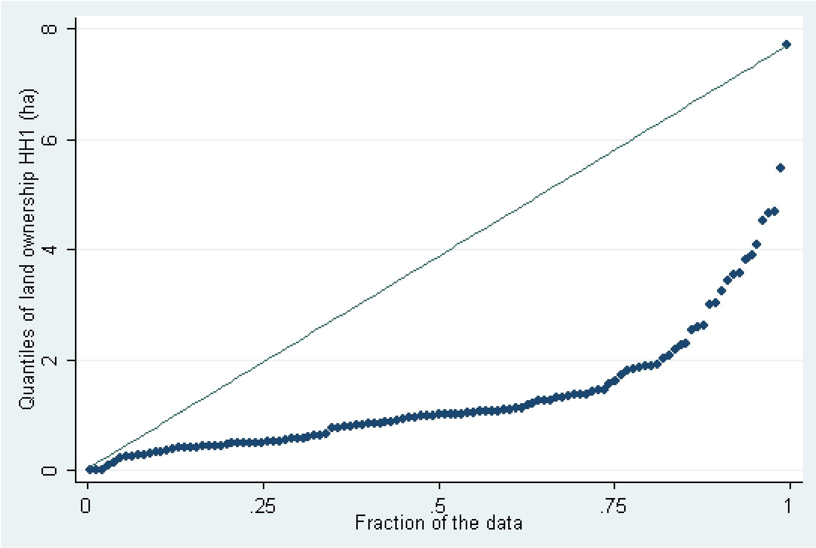


Figure 3: Livestock Quantiles