Does ethnic diversity decrease economic interactions? Evidence from exchange networks in rural Gambia*

Jean-Louis Arcand†
The Graduate Institute

Dany Jaimovich‡
Goethe University Frankfurt

April 29, 2015

Abstract

Using a unique dataset collected in 59 rural Gambian villages, we study how ethnic heterogeneity is related to the structure of four economic exchange networks: land, labor, inputs and credit. We find that different measures of village-level ethnic fragmentation are mostly uncorrelated with network structure. At a more disaggregated level, household heads belonging to ethnic minorities are not less central than those from the predominant ethnicity in any of the networks and, at the dyadic level, the fact that two households share ethnicity is not an economically significant predictor of link formation. Our results indicate that, in the particular setting of our study, the structure of the exchange networks is better defined by other variables than ethnicity, and that ethnic heterogeneity is unlikely to be a driver for sub-optimal economic exchanges. The network structure of our data allow us to include fixed effects at different levels as well as to precisely measure kinship ties, a confounding variable often omitted in previous studies.

We argue that our findings can be interpreted in a causal way as the current distribution of ethnic groups in rural Gambia is largely influenced by specific historical features of the British colonial administration. The colonial authorities divided the territory in districts which closely followed the division of the small pre-colonial Mandingo kingdoms, relying on the appointment of local authorities in most administrative positions. The reason for the establishment of this particular administrative system was related to the pacification process after a series of religious conflicts during the second half of the nineteenth century known as the Soninke-Marabout wars. One important characteristic of the Marabout followers was their multiethnical composition, which formed the basis of the ethnic distribution in each administrative unit. This ethnic diversity was preserved during a pacification process that limited the mobility between districts.

Keywords: West Africa, Social Networks, Ethnic Fragmentation.
JEL codes: C31, 012, Z13.

*We thank Simon Hess, Raymond Jatta, Rajesh Ramachandran, Matthias Schündeln, Chutima Tontarawongsa, Joachim von Braun, Donald R. Wright, and the participants of the presentations organized by the AEL (Bonn), University of Passau, and Goethe University (Frankfurt) for their thoughtful suggestions. Sophie Kazmierczak and Iulia Moldovan provided excellent research assistance.

†Economics Unit, Graduate Institute of International and Development Studies, Pavillion Rigot, Avenue de la Paix 11A , 1202 Genève. E-mail: jean-louis.arcand@graduateinstitute.ch

‡Department of Applied Econometrics and International Economic Policy, Goethe University Frankfurt. RuW Postbox 46, Grüneburgplatz 1, D-60323 Frankfurt am Main. E-mail: jaimovich@econ.uni-frankfurt.de

1 Introduction

Is ethnic diversity detrimental to economic development? The popular answer, especially in the setting of developing countries, tends to be positive. For instance, several studies relate ethnic fragmentation and polarization with warfare and conflicts (see Kanbur *et al.*, 2011, for a recent review). In a particularly influential study, Easterly and Levine (1997) classify ethnic fragmentation as "Africa's growth tragedy" and a voluminous literature has followed them, though their main results are questioned by Arcand *et al.* (2000) and others. One of the arguments to support this view is that different ethnicities mistrust each other and are less likely to create links within social and economic networks: "*lower trust between diverse ethnic groups makes it difficult to form the social networks (social capital) that promote growth by disseminating advanced technology and economically useful knowledge.*" (Easterly, 2001, p.689).

When the effects of ethnic heterogeneity are analyzed using a networks perspective, studies focusing on rural Sub-Saharan Africa have produced results that are by no means clear-cut. Grimard (1997) finds that Ivorian households tend to partially risk-share within the same ethnic group and Arcand and Fafchamps (2012) find that households of the same ethnicity are more likely to belong to the same community based organizations in Burkina Faso and Senegal. While these results seem to provide support to Easterly's aforementioned hypothesis, other findings suggest the opposite. For instance, Hoddinott *et al.* (2009) point out that households belonging to ethnic minorities do not have smaller mutual assistance networks in labor-sharing groups in rural Ethiopia, and Fafchamps (2003) finds no evidence that ethnicity is an important predictor for trust among traders in market towns of Benin, Malawi and Madagascar (when other network characteristics are controlled for).

We aim to contribute to the recent empirical literature concerning the study of social and economic networks in rural economies, particularly focusing on the role of ethnic diversity. We use very detailed data about social and economic interactions collected in 59 Gambian villages to shed light on the following questions: Do villages with different levels of ethnic heterogeneity differ in the structure of their exchange networks? Are households that belong to ethnic minorities segregated in these networks? Does the probability to establish an exchange link increase only because agents are from the same ethnic group? Our findings tend to provide negative answers: Ethnic composition seems to play a limited role in defining the network of exchanges and people from different ethnic groups are not less likely to engage in economic exchanges.

The Gambian context is interesting: a small and geographically homogeneous country where the ethnic composition of rural villages is a mix of various ethno-linguistic groups.[1] Within village ethnic heterogeneity is not a common factor in West Africa, and in the Gambia case is largely related to historical factors. Differently to the case of other territories, the British colonial administration system preserved the diversity existent in

[1]To our knowledge, the only study that has analyzed network related effects of ethnicity in The Gambia is Gajigo and Foltz (2010), which concentrates particularly on the groups of the Serahules, showing that households of this ethnic group are more likely to provide credits among them, a fact that, jointly with the high density of Serahules' geographical distribution, explains the entrepreneurial success of that group. These conclusions are drawn from household survey data and therefore the network interpretation is very limited.

pre-colonial times as part of a pacification process. Recent events like migration and armed conflicts have played a limited role in shaping the current distribution of ethnic groups in rural areas. We argue that this particular feature of the Gambian context make it feasible to interpret the effects of ethnic diversity on economic activities in a causal way. Moreover, the network structure of our data allow us to include fixed effects at different levels as well as to precisely measure kinship ties, a confounding variable often omitted in previous studies.

The present study contributes to the literature in various ways. Firstly, the characteristics of the data allow us to analyze the relation of ethnic diversity and the structure of the various networks of economic exchanges: land, labor, inputs and credit. Secondly, given the availability of information for a large number of villages we can simultaneously analyze the networks at different levels: village (macrostructure), household (mesostructure) and link (microstructure). This allows us to control for both observed and unobserved village characteristics using village fixed effects in the last two levels of analysis, which further strengthen our identification strategy based on the historical formation of the distribution of ethnic groups in the country. Another contribution is that, unlike most previous studies, our network data has detailed information about kinship ties, and therefore we can distinguish between economic exchanges with family members from those with other members of the same ethnic group. In fact, we show that if kinship ties are not considered in the analysis the estimated coefficients are biased towards a negative effect of ethnicity in economic exchanges.[2] Finally, given that we have an almost complete census of links in each village, our estimates are not biased as a result of the selection of a sub-sample of village respondents (Chandrasekhar and Lewis, 2011).

Our main results show that, at the village-level, there is no significant relation of an index of ethnic fragmentation with the network structure of land, labor, input and credit exchanges. This is also the case when an index of polarization is taken as measure of ethnic diversity. If historical data about ethnic fragmentation is considered instead, we find that, if anything, there are positive long-term effects of diversity on economic activity. This is also the case when the historical data is used as an instrumental variable for the estimation of the ethnic fragmentation index. In the analysis at the household-level, our results provide evidence that household heads belonging to ethnic minorities are not less central than those from the predominant ethnicity in any of the networks, and at the dyad-level the prediction of a link creation is barely influenced by shared ethnicity, with estimated coefficients which are either statistically or economically insignificant. Therefore, taking results at different levels, we find limited evidence that ethnic heterogeneity plays a role in the structure of the analyzed economic networks.

The rest of the paper is organized as follows. The next section describes the main ethnic groups present in The Gambia and how their distribution in the country is largely related to pre-colonial circumstances. Section 3 describes details of the data collection and presents descriptive statistics and definitions for the measures of ethnic diversity and the network related variables. Section 4 presents the main empirical analysis. A final

[2]The lack of accurate kinship ties data is a shortcoming of many previous studies. For instance, in their analysis of risk-sharing networks in The Philippines Fafchamps and Gubert (2007) speculate that kinship is a fundamental factor for network formation, but they only have information about geographical distance as a proxy for it.

section concludes.

2 Ethnics groups and their distribution

The objectives of this section are twofold. First, we provide a description of the main ethnic groups of The Gambia and their historical interactions. This is relevant to support the interpretation of our main result, namely that ethnic diversity does not significantly affect economic interactions. Secondly, we show the factors that have shaped the current distribution of ethnic groups in rural Gambia. It is found that particular aspects of the period of consolidation of the British colonial rule play a major role. This fact is relevant as part of our identification strategy relies on the assumption that ethnic heterogeneity is largely determined by historic factors and then exogenous to current interactions in the networks of economic exchanges. If this is not the case, unobserved factors related to ethnic diversity may be driving our results instead.

2.1 Ethnics groups in The Gambia

The sample of 2,689 households in our database reflects the ethnic diversity in The Gambia. 50% of the respondents are Mandinka, 23% Fula, 10% Wollof, 8% Jola, 5% Serer, and 2% Serahule, with the rest either belonging to local ethnic minorities or being non-Gambian (the latter represent 3% of the sample). Compared with the data for rural areas of the 2003 Census for The Gambia (Table 1), the sample overrepresents Mandinkas and underrepresents Serahules.

These groups are genetically diverse and have populated the current area of The Gambia in different times (Allsopp *et al.*, 1992). One way to understand the salience of the differences among the different ethnic groups is the distance of their languages in the linguistic tree. Figure 1 shows the language families in The Gambia and its different branches. The languages tend to be distant from each other, with the only exception of Wollof and Fula.

The Mandinka, also known as Mandingos, arrived over seven hundred years ago with the expansion of the Malian empire. Since British colonial times, they have become mostly farmers growing rice and peanuts. The Fula is an ethnicity widely dispersed in West Africa, further believed to be of North or East African origin. The Fulas have preserved a high degree of ethnic identity, historically related to nomadic pastoral activities and nowadays involved in cattle ownership. The Wolofs are believed to have been forced southward by the Berber expansion even earlier than the Mandingo (about 1,300 years ago) and are situated along the north bank of the Gambian river. Whereas most of the Wolofs are farmers, those living in the area of Banjul have been known for business activities as well as civil service duties (Mwakikagile, 2010). The Serahule are believed to be part of the Soninke people, the founders of the Ghana Empire. They have always primarily been considered merchants since the times of the trans-Saharan trade (Gajigo and Foltz, 2010). Little is known about the history of the Jolas, that are thought of as one of the only descendants of the original inhabitants of the region. Most of the members of this ethnic group live south of the Gambia river and in the Casamance region, cultivating rice. They have been resistant to change and to influences by other ethnic groups and have thus also been less receptive to the spread of the Islam across the region (Quinn,

1972).

Despite the ethnic diversity in the country, several observers have noticed a remarkably harmonious cohabitation of these different groups compared to the rest of the African countries (Mwakikagile, 2010; Wright, 2004; Thomson, 2012). Among the main explanations for this fact are: (i) most people in The Gambia share a common belief -a syncretic variant of the Islam- that serves as a unifying factor but also has been subject of political manipulation to restrain, potentially ethnic-based, opposition movements (Darboe, 2004);[3] (ii) a common cultural influence from the ancient empires and kingdoms, with most of the groups developing a similar social hierarchical structures in the past (where inter-ethnic marriages within a social class were common) and currently sharing Mandinka as their *lingua franca* (Barry, 1998; Wright, 1999; Thomson, 2011); (iii) the influence of the British colonial system, where chief districts were defined by territory and not ethnic groups as was typically done in other African colonies (Thomson, 2012).

2.2 The distribution of the ethnic groups

In an extensive review of the related literature, a recent study by Ahlerup and Olsson (2012) list the factors that have been studied as determinants of ethnic diversity. A first group of factors, classified as the "evolutionary view" are mainly related to geographical characteristics and the pattern of settlement of the first groups of humans. For instance, Michalopoulos (2012) shows that geographic variability is a fundamental determinant of ethnic diversity worldwide and Ahlerup and Olsson (2012) that the date of the first human settlement is also a very strong predictor. On the other hand, the "constructivist view" assigns a greater importance to more recent factors such as the effects of Western colonialism, the state building process, and recent migration. In what follows we describe how each of these factors has influenced the distribution of the different ethnic groups in The Gambia. We conclude that the main factors that have driven the ethnic diversity in rural Gambia are related to events preceding independence.

2.2.1 Historical patterns of settlement

Knowledge of West African prehistory is rather limited and it was only with the establishment of the trans-Saharan trade routes that written history became available.[4] The first known dominant culture of the area was the Ghana Empire in the middle of the eleventh century, which was replaced by the Mali (Manding) Empire roughly two hundred years later (Faal, 1997). After the decline of the Mali Empire in the fifteenth century, a number of Mandingo kingdoms clustered along the Gambia river banks, some of them still existing during the British colonial rule (Barry, 1998). The history of coexistence of the different ethnic groups under these kingdoms was not always peaceful, since Mandingo domination seems to have required a forceful expansion as well as exploitation of neighboring ethnic

[3] In our sample 96% of the respondents declared to be Muslims, the same proportion as in the 2003 census in rural areas. In the census 5% of the people declared to be Christian, but they are mainly located in urban areas.

[4] Many aspects of the history of The Gambia and its villages have been preserved by the Griots (traditional story tellers) that have narrated and passed stories from one generation to the other for centuries.

groupings resulting for instance in armed disputes between Mandingo and Wolof over the principal Gambian port Kau-ur in the nineteenth century as well as the exploitation of Fulas through the payment of taxes and delivery of services to the Mandingo aristocracy (Quinn, 1972). Nevertheless, inter-marriages and the protection given to people from other ethnicities that accepted the Mandingo authority created a multiethnic society.

The patterns of population distribution in the Senegambia area are related to its involvement in three historical major trading zones: the trans-Sahara trade to the North, the Sudanese trade along the Niger Bend to the South as well as the forest trading circuits of Sierra Leone. It was through these routes that Islam first arrived in the area, consolidating its influence during the wars against the European slave traders led by religious leaders in the seventeenth century (Barry, 1998). Trade with the Europeans, including an extensive slave trade, was added later as another major factor in defining economic activities and consequently the distribution of ethnic groups. The Portuguese were the first Europeans to establish in the region in the sixteenth century. Later France consolidated its hold on the Senegal area in the seventeenth century while England captured James Island from the Dutch in 1661, starting their rule over the Gambia river area and their first settlement on the West African coast.

2.2.2 The colonial system

It was only in 1889, with the treaty that divided the Senegambia region between the French and the British, that the colonial system was consolidated. The system of administration established by the British in The Gambia will have long-lasting consequences in the distribution of ethnic groups in The Gambia. The colonial authorities divided the territory in districts which closely followed the division of the small pre-colonial Mandingo kingdoms, relying on the appointment of local authorities in most administrative positions (Gailey, 1965; Thomson, 2012).

The reason for the establishment of this particular administrative system was related to the pacification process after the Soninke-Marabout wars. These were a series of conflicts during the second half of the nineteenth century that confronted Mandingo chiefs (Soninke), who were not fully adherent to Islamic practices, with followers of religious leaders that demanded a more important position of the Muslims in the political structure. As a result, the Marabout followers, which included Mandingo Muslims as well as people from other ethnic groups, gained great influence even challenging the control of the colonial rulers, which confronted them and exiled their main leaders. While some of the Soninkes regained part of their power under the British rule, the rival Marabouts remained highly influential and their importance was confirmed by formal appointments from the colonial authorities (Quinn, 1972). One important characteristic of the Marabout followers was their multiethnical composition, which formed the basis of the ethnic distribution in each administrative unit (Quinn, 1972; Wright, 2004).

Mobility between these districts was initially limited as a result of the pacification process and afterward by the agricultural production system implemented by the British authorities. The labor shortage for the production of groundnuts, the main cash crop, was mainly covered with the use of *strange farmers*, seasonal migrants coming mainly

from neighboring French and Portuguese colonies, more than through internal migration (Swindell, 1978). The strange farmers labor force was very large. As described by Jarrett (1949), during the first half of the twentieth century it was estimated in 15,000 to 20,000, as compared with a local population of no more than 200,000.

The idea that a British colonial pacification process restricted mobility and defined the current ethnic distribution is indeed similar to the identification strategy proposed by Miguel and Gugerty (2005) in their study of ethnicity and public goods provision in Western Kenya. While similar to the historical conditions described by them, the Gambian colonial institutions differ from those in other former British territories in Africa.[5] While in many colonies indirect ruling by local chiefs was basically divided by ethnic lines, the fact that the ancient Mandingo kingdoms were in practice ethnically diversified entities, jointly with the multiethnic composition of Marabout followers during the Soninke-Marabout wars, implied that the colonial administrative units in The Gambia were mainly defined by the pre-colonial territorial demarcation. As a consequence, the current distribution of ethnic groups is likely to be similar to the one at the consolidation of the British rule.[6]

2.2.3 The state building process

Previous studies have also mentioned the process of state building as a potential determinant of the distribution of different ethnic groups in Africa. For instance, Miguel (2004) shows how the formation of the state after independence differently affected the interethnic cooperation process in neighboring districts of Kenya and Tanzania. In the case of The Gambia, this process seems to have a limit role, particularly given the absence of (ethnic based) civil conflicts.

The global opinion after independence was that the country was too small to remain independent and that integration with Senegal ought to be considered.[7] In spite of these claims, the first Gambian political leaders managed to develop a sense of national identity across ethnic groups and eventually pursued the ideal of remaining independent (Sallah, 1990). During the presidency of Dawda Jawara, who led the country from independence until 1994, the representation of minority ethnic groups as members of the government and civil service was promoted with the aim of fostering national integration as well as support for Jawara's PPP party (Edie, 2000; Perfect, 2008).[8]

After the military coup led by Yahya Jammeh (of Jola origin) in 1994, who then became the leader of the country until today, the new government has also put great

[5] Lange (2004) shows the large heterogeneity in the type of institutions established by the British colonial authorities.

[6] The historical identification strategy would be invalid if villages in our sample were recently founded. This is not the case, as the an average of 73% of the village inhabitants declared that their family lives in the village for many generations. This is also the of 78% if only respondents that are more than 50 years old are considered.

[7] A failed attempt to establish a Senegambia federation started in 1982, but only lasted for seven years.

[8] Nevertheless, the Mandinka, to which Jawara belonged, remained the most influential group (Saine, 2008).

emphasis on the inclusion of the different ethnicities, as a way to restrain the surge of ethnic-based opposition groups. This search for coexistence of the different ethnic groups is also reflected at the local level in rural organizations, where the constitutions of village based groups tend to recognize and promote the participation of members from different ethnic groups (Thomson, 2012).

2.2.4 Recent migration

Even though ethnically related conflicts have not affected the distribution of the ethnic groups, recent economic migration could have reshaped it. If this were true, it may be the case that migration patterns could be simultaneously related to the structure of the networks and the ethnic diversity of the village, and therefore the results presented in the Section 4 may be biased. Gambia Bureau of Statistics (2007) shows that, according to data from the 2003 census, most of the internal migrants have moved from rural to urban areas. As a consequence, the rural population in the country has gone from 62% in 1990 to 43% in 2010 (World Bank, 2013). Nonetheless, this process seems not to be related to some particular ethnic groups. The last two columns of Table 1 show that, according to data from the national censuses, the share of the different ethnic groups in the villages of our sample barely changes between 1993 and 2003 for the 55 villages where data was available (the difference in shares is not statistically significant). Moreover, these shares are very similar to those of our data (Table 1, column 1).[9]

While on a smaller scale than the urban-rural migration, the migration within rural areas is also a relevant process. According to the data from the 2003 National Census, an average of 10% of the rural village inhabitants were not born in the village and come from other rural areas, another 3% come from urban areas, and less than 1% come from another country (mainly Senegal). Nevertheless, recent migration patterns do not seem to explain ethnic diversity. In the 1,811 rural villages for which data is available in the census 2003, the correlation between the percentage of rural immigrants and the index of ethnic fragmentation is very small (ρ=0.09). In Appendix section A we provide a series of regression analyses which support the idea that recent migration explain only a small portion of the variation of ethnic diversity in rural Gambian villages.

3 The data

3.1 Data collection strategy

The data were collected by the authors, other researchers, and local collaborators in the context of the baseline survey for the impact evaluation at national level of a Community-Driven Development Program (CDDP), conducted between February and May of 2009. 59 Gambian villages[10] with populations between 300 and 1,000 inhabitants, mainly in rural areas (just 3 villages are in semi-urban areas), were randomly selected using area sampling at the ward level, a smaller geographical division that tends to be homogeneous

[9] Despite the fact that we based our definition on the ethnicity of the household heads instead of total village population.

[10] We have excluded one of the villages from the original sample, because all the inhabitants were migrants from Ghana and their ethnicity was not specified (they were only recorded as "non-Gambians").

in geographical but heterogeneous in socio-cultural terms.[11]

The data collection strategy adopted for the present study differs from that of traditional household surveys and instead relies on structured group interviews geared to collect quantitative information. Therefore, village gatherings co-organized with the village chief (*Alkalo*) and other authorities were carried out. In such meetings it was possible to obtain coarse quantitative information -with a particular focus on socio-economic interactions- for almost all household heads in each village.

In the structured group surveys two categories of information were collected. The first section was a standard (and very lean) household questionnaire designed to collect a vector of household characteristics including: economic and demographic indicators, traditional roles in the village, membership in various community based organizations, and other household characteristics. The second section of the survey instrument was specifically designed to understand the economic and social networks in the village using the following questions:

- $LAND$: *Did members of your household lend out or borrow in land from other villagers?*
- $LABOR$: *Did you, or any members of your household, work for other households during the last year?*
- $INPUT$: *Did members of your household lend out to or borrow in any means of production (such as tools or fertilizer) from other households in the last year?*
- $CREDIT$: *Did members of your household lend out to or borrow in money from other household in last year)?*
- $MARRIAGE$: *Have any of your household members married members of other households?*
- $KINSHIP$: *With which households do your household members have direct kinship relationships?*

The information related to our main questions (links in the economic networks) was usually observable for community members and therefore common knowledge and not sensitive information that would not be revealed in public.[12] Moreover, any particular community characteristic which may be correlated with measurement error in the data is not a problem because in most empirical specifications village fixed effects are included. In terms of household characteristics related to the measurement error, in all specifications a rich set of household level controls are included and household fixed effects are included whenever feasible.

[11]The CDDP project targeted the bottom half of the poorest villages in The Gambia, according to a poverty index based on Census 2003 data. When the population limit (between 300 and 1,000 inhabitants) is considered, our sample is representative of roughly 20% of the rural villages in The Gambia.

[12]The only exception may be credit exchanges. During the pilot surveys we were initially reluctant to ask information about credit links, but it was found that villagers, and in particular borrowers, were in general willing to respond to this questions. Still, the analysis regarding this particular network must be taken with a grain of salt. In any case, the main results do not depend on it.

One of the main advantages of this data collection strategy is the fact that kinship relations within the village are very precisely measured. This is the case because the blood ties and marriage relationships are revealed by all the respondents, therefore the network structure of all kinship relationships can be recovered even if not full information is reported by everyone. Below we will show that accurate kinship data is relevant, as otherwise the results related to ethnic diversity may be purely driven by interactions with family members, which are mostly from the same ethnic group.

3.2 Data description

In order to minimize selection problems it was attempted to interview all households in the village, yielding a median coverage rate of 94%. Finally, 2,886 persons were interviewed, but the sample was reduced to 2,689 when incomplete data were removed. Table 2 presents descriptive statistics for the data collected in the general questionnaire, at both village- and household-level. The smallest village has 202 inhabitants and the largest 1,402.[13] The average population of the surveyed villages is 587 inhabitants. Population density, at least when the denominator is the inhabited area, is high, with an average of 6,911 inhabitants per square kilometer. In contrast, agricultural land is usually very abundant. The average amount of agricultural land per active worker is around two hectares, when land usage rights for the year of the survey were considered. Average household size, on the basis of the household data, is 12.7 members.[14] 47% of households declare to be polygamous and, as is to be expected in West Africa, a very small number of household heads are females (6%), generally represented by widows and mainly concentrated in the semi-urban areas on the outskirts of Banjul (the national capital). These villages also accounted for the few non-Muslims in the sample (less than 1%).

The economic conditions in the villages in the sample correspond, by and large, to those of traditional rural societies. There is almost no access to electricity, with an average village-level access rate of 3%. 88% of households have no access to an improved source of water, while 37% lack access to a private toilet. 38% of household dwellings are built mainly with grass. 83% of the respondents declared having no formal education, although a substantial fraction of the villagers received some kind of Koranic education and usually master basic Arabic language skills. For the empirical analysis, these variable will be compounded in a poverty index that is described in Table 2.[15] The main economic activity is related to agriculture (66% of households have this as their main activity) or fisheries (6%). Nevertheless, a Herfindahl index of sectoral heterogeneity shows a significant degree of diversity, driven mainly by the presence of inhabitants working outside the village (25%). Monetary income is very low. The average (self-declared) annual income per capita is 3,514 Gambian Dalasis (which corresponds to $380 in constant 2005 and PPP adjusted dollars from World Development Indicators), and just around 12% of this

[13]The target population was 300 to 1,000 inhabitants, but the data from the 2003 census in which the sampling was based was sometimes inaccurate or village population had changed in the last six year.

[14]While some households appear exceedingly large, respondents were very clear in terms of their definition of a household. The presence of households with more than 50 members (approximately 1% of the sample) is explained both by the polygamous nature of Gambian rural society and the existence of *marabout* households where the household is constituted by a mass of disciples and other followers.

[15]This is the same poverty index used by the World Bank to target the implementation of the CDDP project

income stems from agricultural activities. The average Gini coefficient is 0.34, but it reaches 0.66 in some villages.

More details about the data collection process and a more extended description of the data can be found in Jaimovich (2011) and Jaimovich (2015).

3.3 Measures of ethnic heterogeneity

As already mentioned in the previous section, our sample represents the diversity of ethnic groups in The Gambia (Table 1). One aspect which makes the country an interesting place for the study of the effects of ethnic diversity is the large variation of ethnic composition within villages. In our sample, 21% of the respondents come from a minority ethnic group in their village, and most of them (17%) actually come from villages where their group accounts for no more than 30% of the population.

Two indices will be used to measure ethnic heterogeneity within each village. The most commonly used measure of ethnic diversity in the empirical literature is the index of fractionalization, and particularly the ethno-linguistic fractionalization index which is defined as:

$$ELF = (1 - \sum_{e=1}^{N} s_e^2),$$

where s_e is the share of the ethnic group e in the village. The ELF for the 59 villages in our sample ranges from zero (complete homogeneity) to 0.85, with a mean of 0.29. As shown in Figure 2, the distribution of the ELF has observations along most of the support, even though it is skewed towards more homogeneous villages. This distribution is actually similar to that of all rural villages in The Gambia, as shown in Figure 3 (which considers data from the 2003 national census).

In some of our empirical estimations we will use the ELF calculated from the 1993 census data, ELF_{1993}. This variable has exactly the same mean as the ELF from our data, but reaches a maximum value of 0.98. The difference between the two may be related to the fact that in our data fractionalization is calculated based on the ethnicity of the household head, while in the census all the inhabitants of the village are considered. Despite this difference, the correlation between ELF and ELF_{1993} is very high (0.8), which provides further evidence related for the hypothesis that ethnic diversity does not change significantly over time in rural Gambia.

A different measure of ethnic diversity is proposed by Montalvo and Reynal-Querol (2005). Based on conflict and rent-seeking theories, they develop an index from the family of polarization measures defined as:

$$POL = 4 \sum_{e=1}^{N} \sum_{i \neq e} s_e^2 s_i,$$

which measures the normalized distance of a distribution of ethnic groups from a bimodal distribution. Therefore, POL gets the highest value when there are two groups of the same size, and low values if there is one very predominant group and one or more

small groups. The distribution of POL in our data is bimodal (Figure 2), with modes at small values (homogeneous villages) and values above 0.8 (villages with two large ethnic groups). As it was the case for the ELF, this distribution is closely related to the distribution of POL in all rural Gambia villages (Figure 3).

Both measures of ethnic diversity are highly correlated (0.88), but this is mainly related to villages with low values in both indices. As it can be seen in Figure 4, for high values the relationship between the two variables is actually negative. Interestingly, the relationship between POL and ELF displayed in Figure 4 is similar to the one shown by Montalvo and Reynal-Querol (2005) when data for 138 countries are considered, supporting the idea that the distribution of ethnic diversity measures in The Gambia has large variation. This is further confirmed in the relationship between the indices which considers data from National Census 2003 for all Gambian villages displayed if Figure 5.

3.4 Networks data and measurement

We consider a household as a node i in one of the economic exchange networks m ($LAND$, $LABOR$, $INPUT$, and $CREDIT$). The existence of a link between households i and j in the network m will be measured as a binary variable:

$$\ell_{ij}(m) = 1 \text{ if a link is reported in the data, } \ell_{ij}(m) = 0 \text{ otherwise.}^{16}$$

Given the directed nature of the data, $\ell_{ij}(m)$ is a link from i to j, which implies that the former *lends* m to the latter. If the opposite is true (i borrows from j), then the link will denoted as $\ell_{ji}(m)$.

A basic metric of the embeddedness of a node i in a network m is its *degree*, $d_i(m)$, measured as the number of links involving this particular node. In the case of the data, a distinction must be made depending on the directionality of the link. If the link goes from i to j, then it is counted in the measure of the *out-degree*. Formally:

$$\text{Out-degree: } d_i^{out}(m) = \sum_j \ell_{ij}(m).$$

In the economic networks, the out-degree of i is related to its position as a lender. When the link goes in the other direction, from j to i, it will be counted as part of the *in-degree* of i:

$$\text{In-degree: } d_i^{in}(m) = \sum_j \ell_{ji}(m).$$

For economic networks the in-degree is a characteristic of i as borrower. In order to be able to compare across networks, degree centrality will be expressed as a proportion of the total possible links in each particular network in a village (n):

$$\text{Degree centrality: } C_i(m) = \frac{d_i(m)}{n-1}.$$

[16] A link is recorded in the data if at least one of the two villagers of the dyad mention the existence of a link.

As can be seen in Table 3, the average degree centrality tends to be between 1% and 2% of total possible links, except for $KINSHIP$ with 10%, but the data are very heterogeneous, indicating important differences in the centrality of households.

In the present study, the analysis of network architecture or *macrostructure* is at the village-level. A common measure used in networks analysis at this level is the density of the network, the sum of the degrees of all households in the network m over total possible links:

$$\text{Density: } D(m) = \frac{\sum_{i=1}^{n} d_i(m)}{(n-1)n}.$$

The density can be considered as the probability of forming a link in a purely random network of average degree $d_i(m)$. In our data (Table 3), densities range between 2% and 5% for most networks, with the exception of $KINSHIP$ which has an average density of 14%, confirming the fact that in several villages the inhabitants tend to be close relatives. If ethnic diversity reduces the probability of a link formation, it is expected that villages with higher ELF or POL will have lower density.

An important regularity observed in the networks analysis literature is that an ego's links, instead of being distributed evenly, tend to be concentrated in *local neighborhoods*, creating clusters of egos well connected among themselves but not with the rest of the network ("my friends tend to be friends amongst themselves"). A standard measure used to study this feature is the clustering coefficient, $Cl_i(m)$, which represents the probability that j and k are linked given that i has a link with both:

$$\text{Clustering coefficient: } Cl_i(m) = \frac{\sum_{j\neq i;k\neq j;k\neq i} \ell_{ij}\ell_{ik}\ell_{jk}}{\sum_{j\neq i;k\neq j;k\neq i} \ell_{ij}\ell_{ik}}.$$

The clustering coefficient for a given network m can be calculated as a weighted average of $Cl_i(m)$:

$$Cl(m) = \sum_i Cl_i(m)w_i,$$

where w_i weights the density of each local neighborhood proportional to its size. Table 3 shows that, as it is to be expected, $KINSHIP$ links are very clustered (mean $Cl(KINSHIP)$ is 0.28) but $LAND$ links are not. If households mostly interact with villagers of the same ethnicity, then it must be expected that $Cl(m)$ is higher in more ethnically diverse villages.

In a random network, a standard prediction is that as the probability of forming a link increases the nodes will be grouped in a few components (nodes connected among themselves but not with rest of the network) and eventually a *giant component* that connects most of the elements in the network will emerge. But when the formation of networks is strategic the links can be distributed in very different ways. In particular, we would expect that if ethnic groups just interact among themselves, the final structure of the network will be less *compact*. To explore this possibility, we propose a measure called *index of compactness* that measures the dispersion of the links in different components,

defined as:[17]

$$\text{Index of compactness: } Cmp(m) = \sum_{c=1}^{C} s_c^2,$$

where s_c is the share of nodes in each component as a fraction of the total number of nodes. When the network is a collection of several small components we expect small values of $Cmp(m)$, as is the case for $CREDIT$ in our data (Table 3); if links are concentrated in few big components $Cmp(m)$ take values around 0.2 and 0.5, which is the case of the other economic networks in our data, and when most of the nodes are in a giant component $Cmp(m)$ takes values above 0.5, as in $KINSHIP$.

4 Main results

In this section we use our unique data from Gambian villages to analyze the relationship between ethnic heterogeneity and interactions in four networks of economic exchanges. This empirical specification considers three levels of disaggregation: village (network "macrostructure"), household, and dyad (link).

4.1 Macrostructure: network architecture

We will start our empirical analysis with a description of the village characteristics associated with network architecture, with particular focus on the effects associated with measures of village-level ethnic diversity. The latter is measured using the two indices described in section 3.3, namely the ethnic fractionalization index (ELF) and the polarization index (POL).

The main empirical specification will be as follows:

$$y_{mv} = \alpha_{ward} + \alpha_m + \alpha_{ethnic} + X_v\beta + ED_v\beta_{ethdiv} + e_{mv} \quad (1)$$

where y_{mv} is one of the three village-level network characteristics described below: density, clustering and compactness and X_v is a vector of village-level control variables. The coefficient of interest is β_{ethdiv}, which captures the effect of the measures of ethnic diversity ED_v, either ELF_v or POL_v. A set of fixed effects at different levels is considered. Taking advantage of the stratified nature of the sampling scheme, ward-specific effects (α_{ward}) that control for several geographical characteristics, such as distance to the capital and other important population centers or agro-climatic characteristics. Given the possibility that results related to ethnic diversity can be driven by the characteristics of the predominant ethnicity, in all specifications a dummy identifying the largest ethnic group in the village (α_{ethnic}) is included. In all the specifications where data are pooled for more than one network, network-specific fixed effects are included as well (α_m).

In our main specifications, equation 1 is estimated using OLS. Nevertheless, since the left hand side variable (the network measures) is always a proportion, we check the robustness of our results to estimations with the quasi-MLE *fractional logit* procedure

[17] This index is similar to the *component size heterogeneity* index using by sociologists.

proposed by Papke and Wooldridge (1996).[18]

Table 4 presents the results when the ELF is used as the measure of ethnic diversity and the estimation is over the pooled set of networks, as well as when dividing the sample into *economic networks* ($LAND$, $LABOR$, $INPUTS$ and $CREDIT$) and *family networks* ($KINSHIP$ and $MARRIAGE$). The upper panel of table shows the ward-level fixed effect estimation when no control variables are included. The explanatory power of the ELF variable is very small, given very low within R^2 values, and only the coefficients for clustering in economic networks and compactness in family networks are statistically significant. When control variables are included (lower panel of Table 4), all the coefficients associated to ELF are very small and far from being statistically significant (all t-statistics are lower than one). When the sample is divided, the coefficients associated with the economic networks tend to be much smaller than those for the family networks. As such, no evidence of any effects of ethnic fragmentation is found in this specification.

Table 4 shows that other variables have explanatory power for the network variables (only those for which statistically significant results were found are displayed). Density and compactness are decreasing in village population and have a positive correlation with average household size in some of the specifications (larger households have more members who can potentially interact with other households in the village). Income inequality (captured by a Gini index for self-declared household income) is positively related to both density and compactness which also the case for the poverty indicators, the poverty index and the proportion of grass huts over corrugated hut's roofs. In the second to last row of coefficients of Table 4 it can be seen that economic networks density and clustering increase with the percentage of migrants in the village (a variable taken from the 2003 National Census).[19].

In the regression where only economic networks are considered, we include the variable *% relatives*, which is measured by the density of the $KINSHIP$ network in each village. The last row of Table 4 shows that villages where inhabitants have more kin ties economic networks are more dense and compact (columns 2 and 8). This results is relevant, as in the more disaggregated estimations presented below it will be shown that ethnicity can be confounded with kinship when information for the latter is not available.

In Table 5 the results of the separate estimations by network are displayed. A more parsimonious model is used, given the reduced number of degrees of freedom in this specification with only 59 observations (details at the bottom of the table). Both the results without and with control variables (which are fairly similar) tend to confirm the main previous finding, with an ELF index which has a low explanatory power and is mostly not statistically significant. In the specification of the lower panel there are two exceptions, as density and compactness in the land network increase with higher values of ELF. This estimated effect is large. In the case of compactness, one standard deviation

[18]In fractional logit estimations we follow Cameron and Windmeijer (1997) in reporting a goodness of fit measure based on the deviance.

[19]We have previously shown that migration is unrelated to ethnic diversity when all rural Gambian villages are considered (Appendix A). Including the percentage of migrants in the village as a control variable further deals with the concern that the estimated β_{ethdiv} coefficient may be biased due to the effects of recent migration.

increase in ELF implies a change of 0.114 in the compactness of the land network, or around a 50% increase from its mean value. As for the results for the control variables, they are similar to those in the pooled networks model.[20]

When POL is used as the measure of ethnic diversity the results are equivalent to those obtained with ELF, as it can be seen in the upper panel of Table 6. POL is never a significant determinant of village networks structure, except for the already discussed positive and significant effect for compactness in the land $LAND$.

As discussed above, ethnic diversity in rural Gambia is mainly determined by historical factors, and for this reason we argue that the observed lack of effect of ethnic diversity may actually be interpreted in a causal way. Notwithstanding, we cannot rule out that recent economic conditions in the village may have unobserved effects on fertility decisions, intra-household migration, or other household decision processes that can bias the estimates. In order to deal with this potential problem, we check if the main results hold when historical data on ethnic diversity is used instead. The oldest available data for the villages in our sample come from the National Census 1993 (data for four villages were not available), where information for all inhabitants of the village is available and therefore the corresponding fractionalization index (ELF_{1993}) can be built not only with the ethnicity of the household head, as it the case in our data. The results of using the historical index in the estimation are displayed in the mid-panel of Table 6. The coefficients for ELF_{1993} are now larger, and positive and significant for density and compactness of the pooled economic networks (column 2). Nonetheless, this result is still similar to the one obtained for ELF, as the regressions by network indicate that the effect is mostly driven by $LAND$ (columns 3 to 8).

As shown in section 3, ELF_{1993} is highly correlated with the ELF index from the 2009 data (rho = 0.78). If the historical index only influences current village economic interactions through its effect on the current fractionalization index, then it can be used as an instrumental variable to deal with the potential reverse causality problem in the estimation of the β_{ethdiv} coefficient, as well as with attenuation bias related to measurement error of the ethnic diversity variables. While this exclusion restriction cannot be directly tested, Appendix Table B.2 shows suggestive evidence to support it, as there are no significant differences at the 5% level in the t-tests of difference in the mean value of thirteen observable characteristics of villages with high and low values of ELF_{1993} (as defined by its median).

The results of the instrumental variable estimation are presented in the lower panel of Table 6. As we expect that the exclusion restriction is particularly likely to hold in the case of the economic networks and the instruments do not vary at the village-level, we only present results for each separate economic network. The first stage F-test of the excluded instruments confirms the strong correlation of the historical and the 2009 fractionalization indices (F-statistic value is around 20). The coefficients of the instrumental

[20]The results for family networks, not shown in Table 5, show the expected positive cluster effects of ethnic diversity on kinship and also provide some evidence for the existence of inter-ethnic marriages, given high values of the ELF are related to a lower clustering coefficient, as well as density and compactness, in $MARRIAGE$

variable estimation are higher than those from OLS,[21] but the previous results remain, with the coefficients for $LAND$ density and compactness positive and statistically significant at 1% level. The fact that the OLS coefficients appeared to be downward biased may be due to attenuation bias given measurement errors, for example because in our data we observe only the ethnicity of the household head instead of all members in the household.

Overall, the results in this section provide no evidence of a decrease in economic interactions related to ethnic diversity in the village. If anything, the estimates that consider historical ethnic diversity data show positive effects on current economic exchanges for the land network.[22]

4.2 Ethnic minorities and household centrality

In this section the variables associated with household network degree centrality are studied. In particular, we are interested in testing if households that belong to an ethnic minority in the village are less central, as it would be predicted by the hypothesis that ethnic diversity reduces economic exchanges. On the other hand if a positive or no effect is found, the results of the last section will be further confirmed. One advantage of the household-level estimations that we will present in this section is that village fixed effects can be included in the estimations, therefore any community-level characteristics that affect simultaneously ethnic diversity and economic exchanges are controlled for. Still, household-level unobserved heterogeneity may be a concern for a causal interpretation of the results.

The main model to be estimated is given by:

$$C_{iv}(m) = \alpha_v + \alpha_{ethnic_{iv}} + X_{iv}\beta + Ethmin_{iv}\delta + \mu_{iv} \quad (2)$$

where village-level fixed effects (α_v) are included, as well as a dummy indicating the household head's ethnic group ($\alpha_{ethnic_{iv}}$). The vector X_{iv} contains various household-level characteristics. The coefficient of interest is δ, which captures the relationship between household centrality in network m ($C_{iv}(m)$) and a variable which indicates if the household head belongs to an ethnic minority ($Ethmin_{iv}$).

Table 7 presents the results of the estimation of equation 2. In this specification $Ethmin_{iv}$ is introduced as a dummy variable which takes value one if the household head belongs to an ethnic minority within the village and zero otherwise. Ethnic minorities are more active as borrowers of land and credit lenders. These statically significant effects are large, as they represent an increase of 50% from the mean value of the dependent variable. In all the other networks, we find no significant relationship between household degree centrality and the ethnicity of the household head. These results suggest that minorities are not particularly less active or discriminated against in any of the networks

[21] The results for the OLS estimation in the sub-sample for which data in 1993 was available are very similar to those in the full sample. The results for the sub-sample are presented in Appendix Table B.3.

[22] When a different functional form is considered, the fractional logit estimates presented in Appendix Table B.1 also show positive effects of ELF on networks characteristics.

and, quite the opposite, they can be more active in some economic activities.[23]

As for the control variables, the percentage of relatives in the village always increases centrality (except for land borrowers). This is also the case for household size in most specifications. Older household heads are less likely to lend labor and borrow inputs and household heads with some level of formal education are less likely to be borrowers of land and inputs. In terms of monetary income, richer households are indeed more likely to lend money. Cash crop producers are more likely to borrow land and labor, but also to lend inputs. Unfortunately, we do not have information regarding the migration status of the household head, but we can control for, the expectedly related variables, number of households emigrants and a dummy taking value one if the household receive remittances.

To further explore the robustness of the results, in the upper panel of Table 8 $Ethmin_{iv}$ is divided into *absolute minorities*, whose household heads belong to minorities that represent less than 30% of the population of the village and *relative minorities*, who belong to a minority that represents an ethnicity that is less than a half of the total village population but more than 30%. There are 461 households in the first group and 88 in the second. The results are similar to the ones in the main specification, and mostly driven by the effect of the *absolute minorities*. In the case of minorities with more then 30% of the population, there is a lower probability of lending land. The main results are also confirmed in the last panel of Table 8, where the proportion of household head's ethnic group in the total village's population is used instead of a dummy variable. Large ethnic groups are found to be less active in inputs and credit lending (columns 5 and 7).

Even though households belonging to ethnic minorities do not have less economic interactions on average, it still may be possible that some particular types of these households do. For instance, it may happen that the poorer or the less educated among the ethnic minorities are those who participate less in the economic life of the village. In order to explore this possibility, in Appendix Table B.5 we estimate the heterogeneous effect of $Ethmin_{iv}$ by adding interaction terms with the other household characteristics. Only two of the 40 interaction coefficients displayed in this table are statistically significant, so we can reject the hypothesis of heterogeneous effect of $Ethmin_{iv}$.

So far, we have shown that both in estimations at village- and household-level ethnic diversity does not seem to reduce economic interactions. Why are our findings different from those of the previous literature showing a negative effect? While a first answer is related to the particular characteristics of Gambian communities described in section 2, another possible explanation is coming from the fact that our data contain information usually not available in a typical household survey. In particular, from the network data we know each household's kinship relations in the village, from where we have calculated the variable *% of relatives in the village*. In appendix Table B.6 we explore the effects of not including this variable in the estimation. As compared to the results in Table 8, almost all coefficients are reduced, and actually in some cases a negative and significant effect of $Ethmin_{iv}$ is found, as for labor borrowers in the middle panel. We take this as

[23] As it was the case of equation 1, the dependent variable, be it the in-degree (borrowers/receivers) or the out-degree (lenders/senders) centrality, is expressed as a proportion over total potential links in the network. Appendix Table B.4, shows that when the fractional logit estimation is implemented, the results can be interpreted in a similar fashion as in the OLS estimation.

suggestive evidence that the lack of information on family ties can bias the estimation of the effects of ethnic diversity.

4.3 Ethnic groups and link formation

In the empirical specifications presented above, the network characteristics have been aggregated at the village- and household-level, therefore missing part of the richness of these detailed data. In order to further analyze the effects of ethnic diversity on economic interactions, in this section we present a model where the formation of a link $\ell_{ij}(m)$ with a fellow villager is estimated in a dyadic regression framework. One of the main advantages of this model is that household-level fixed effects can be added, therefore reducing the concern of biased estimated coefficients related to unobserved heterogeneity. Still, some dyadic-level characteristics may remain uncontrolled for, and therefore the estimates must be interpreted with this concern in mind.

The dyadic model has the following structure:

$$\ell_{ijv}(m) = \alpha_i + w_{ijv}\beta_{dyad} + (X_{iv} + X_{jv})\beta_{sum} + |X_{iv} - X_{jv}|\beta_{dif} + \epsilon_{ijv} \quad (3)$$

where the vectors X_{iv} and X_{jv} are the socio-economic characteristics, roles in the village, and existence of external links in each network for both potential exchange partners in a dyad (i and j). Given that the relevant dependent variable is the existence of a link between two households, the undirected formation of links is studied, which implies that $\ell^m_{ijv} = \ell^m_{jiv}$. To preserve symmetry on the right-hand-side, we follow Fafchamps and Gubert (2007) by specifying three types of regressors: the coefficient β_{dif} is associated with the absolute difference in attributes between i and j; β_{sum} is associated with the sum of the attributes of the members of the dyad; and β_{dyad} is the parameter associated with the variable w_{ijv} that indicates common characteristics of i and j. The particular variable of interest in our case belongs to the latter set of variables, and is a dummy that captures whether household heads in each dyad are members of the same ethnic group. The disturbance terms are allowed to be correlated across observations involving the same individual using the two-dimensional clustering methodology described in Cameron *et al.* (2011).

The results of the OLS estimation of Equation 3 are shown in Table 9, where the coefficients for the dummy *same ethnic group* and a selected group of control variables are displayed. In the first column the probability of a link in any of the four economics links is considered ("ALL NETWORKS"). Households belonging to the same ethnicity has a positive effect for link formation at the 10% level. Nevertheless, the magnitude of this effect is very small (0.5% increase in the probability of a link) as compared, for instance, with the effect of a direct kinship relation (consanguinity) between household heads (16%) or kinship relations through marriage (6%). In the other four columns networks are studies separately. Households from the same ethnic group are found to be less likely to have a land link but more likely to form a labor and input link. Again, the magnitude of these coefficients is very small compared with the effect of kinship relations.

In terms of the other variables presented in Table 9, a higher sum of the household size increases the probability of link formation. The coefficient of the sum of self-declared

income *per capita* indicates that richer households are more likely to exchange land and labor between them, while the coefficient of the difference implies the opposite for exchanges between rich and poor households. More educated households are less likely to exchange inputs among themselves and the village chief (Alkalo) is more likely to participate in all economic exchange in the village.

In the previous section it was already shown that the lack of relevance of ethnicity in the level of economic interactions in our estimations may be related to the fact that we are able to properly control for interactions with direct relatives. In Table 10 we report the estimated coefficients for a dummy capturing whether both members in the dyad are from the same ethnic group, omitting the information about kinship (the two dummies capturing blood ties and kinship through marriage). It can be seen that in this case all the coefficients importantly increase in magnitude. The estimated increase in the probability of a link formation in any of the economic networks related to ethnicity is 2.4% and for a labor link 1.4%. Additionally, the effect on the formation of a credit link is now positive and statistically significant. Therefore, as it was the case in the household-level regressions, in the dyadic specifications we again find evidence that omitting family ties information implies an upward bias in the estimation of the effect of ethnic diversity.

5 Conclusions

We have investigated the structure of four economic networks at different levels of disaggregation in a database collected in rural Gambia, with a sample of 59 villages, 2,689 households and 70,007 (potential) links. We use these particularly rich data to contribute to the analysis of the economic development of these communities using a network perspective, with a particular focus on the role of ethnic diversity.

Previous studies, mostly in sub-Saharan Africa, have found evidence of negative effects of ethnic diversity on social and economic interactions. Contrary to this, we do not find a significant relationship between an ethnic fractionalization index and the village-level structure of land, labor, input and credit exchange networks. This is also the case when an index of polarization is taken as the measure of ethnic diversity. If historical data about ethnic fragmentation is considered instead, we find that, if anything, there are positive long-term effects of diversity on economic activity. In the analysis at the household-level, our results provide evidence that household heads belonging to ethnic minorities are not less central than those from the predominant ethnicity in any of the networks, and at the dyad-level the prediction of a link creation is barely influenced by shared ethnicity.

To sum up, we find little conclusive evidence that ethnic diversity plays a role in shaping the structure of economic networks or that people from ethnic minorities are less likely to engage in economic exchanges.

Our identification strategy relies on the idea that the patterns of settlement of the different ethnic groups are determined by particular features of the British colonial administration system and the state building process in the Gambian context. Indeed,

we use historical data on the distribution of ethnic groups whenever possible. Another fact which further supports our findings is that, given the characteristics of our network data, we are able to show that the results hold at different levels of specification and are robust to the inclusion of both village- and household-level fixed effects. Still, improvements in the identification strategy that can provide further evidence of the effects of ethnic heterogeneity on economic interactions remain a fruitful area for future research.

Our results need to be understood in the particular context of rural Gambia, where the ethnic composition of villages is often heterogeneous, but where, despite the traditions preserved by each ethnicity, religion and culture are common to most villagers. In that sense, we do not expect that the same results will be found in other places of rural Africa where ethnic identification is stronger and religious belief differ. Furthermore, we have shown that the lack of correlation of ethnic diversity with economic activities holds only when we control for our precisely measured variable capturing the presence of relatives in the village, raising the possibility that previous studies actually present biased estimates of the influence of ethnicity. As a consequence, one of the conclusions of our study is that future research on this topic must consider that ethnic ties may only be a proxy for family interactions when the latter variable is omitted or imperfectly measured.

Does ethnic diversity decrease economic interactions? Our findings point to a negative answer in the context of rural Gambia. Because of this, we expect that our study contributes to the development of a methodological framework to expand the current body of evidence about the effects of ethnic diversity on social and economic exchanges using detailed network data.

References

Ahlerup, P. and Olsson, O. (2012). The roots of ethnic diversity. *Journal of Economic Growth*, **17** (2), 71–102.

Allsopp, C., Harding, R., Taylor, C., Bunce, M., Kwiatkowski, D., Anstey, N., Brewster, D., McMichael, A., Greenwood, B. and Hill, A. (1992). Interethnic genetic differentiation in Africa: HLA class I antigens in The Gambia. *American journal of human genetics*, **50** (2), 411.

Arcand, J.-L. and Fafchamps, M. (2012). Matching in Community-Based Organisations. *Journal of Development Economics*, **98** (2), 203–219.

—, Guillaumont, P. and Jeanneney, S. G. (2000). How to make a tragedy: on the alleged effect of ethnicity on growth. *Journal of International Development*, **12** (7), 925–938.

Barry, B. (1998). *Senegambia and the Atlantic slave trade*. Cambridge University Press.

Cameron, C., Gelbach, J. and Miller, D. (2011). Robust inference with multiway clustering. *Journal of Business and Economic Statistics*, **29** (2), 238–249.

— and Windmeijer, F. (1997). An R-squared measure of goodness of fit for some common nonlinear regression models. *Journal of Econometrics*, **77** (2), 329–342.

Chandrasekhar, A. and Lewis, R. (2011). *Econometrics of sampled networks*. Unpublished manuscript, Massachusetts Institute of Technology.

Darboe, M. (2004). Islamism in West Africa: Gambia. *African studies review*, pp. 73–82.

Easterly, W. (2001). Can institutions resolve ethnic conflict? *Economic Development and Cultural Change*, **49** (4), 687–706.

Easterly, W. R. and Levine, R. (1997). Africa's growth tragedy: Policies and ethnic divisions. *Quarterly Journal of Economics*, **112** (4), 1203–1250.

Edie, C. J. (2000). Democracy in The Gambia: Past, present and prospects for the future. *Africa Development*, **25** (3), 161–198.

Faal, D. (1997). *History of the Gambia, AD 1000 to 1965*. Serekunda, The Gambia: Edward Francis Small Printing Press.

Fafchamps, M. (2003). Ethnicity and networks in african trade. *The B.E. Journal of Economic Analysis & Policy*, **2** (1).

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

Gailey, H. (1965). *Handbook of the economics of giving, altruism and reciprocity: Foundations*. New York: Praeger.

Gajigo, O. and Foltz, J. D. (2010). *Ethnic Networks and Enterprise Credit: The Serahules of The Gambia*. 2010 annual meeting, July 25-27, 2010, Denver, Colorado, Agricultural and Applied Economics Association.

Gambia Bureau of Statistics (2007). *National Migration Analysis*. Tech. rep., Analysis of the population and housing census 2003.

Grimard, F. (1997). Household consumption smoothing through ethnic ties: evidence from Cote d'Ivoire. *Journal of Development Economics*, **53** (2), 391–422.

Hoddinott, J., Dercon, S. and Krishnan, P. (2009). Networks and informal mutual support in 15 Ethiopian villages. In C. P. Johann F. Kirsten, Andrew R. Dorward and N. Vink (eds.), *Institutional Economics Perspectives on African Agricultural Development*, International Food Policy Research Institute.

Jaimovich, D. (2011). *Macrostructure and microstructure: Evidence from overlapping village networks in The Gambia*. MPRA Paper 38932.

— (2015). Missing links, missing markets: Evidence of the transformation process in the economic networks of Gambian villages. *World Development*, **66**, 645–664.

Jarrett, H. R. (1949). The strange farmers of the Gambia. *Geographical Review*, **39** (04), 649–657.

Kanbur, R., Rajaram, P. K. and Varshney, A. (2011). Ethnic diversity and ethnic strife. An interdisciplinary perspective. *World Development*, **39** (2), 147–158.

Lange, M. K. (2004). British colonial legacies and political development. *World Development*, **32** (6), 905–922.

Lewis, P., Simons, G. F. and Fennig, C. D. (2014). *Ethnologue: Languages of the world*. Dallas, TX: SIL International, 17th edn.

Michalopoulos, S. (2012). The origins of ethnolinguistic diversity. *American Economic Review*, **102** (4), 1508–39.

Miguel, E. (2004). Tribe or nation? nation building and public goods in Kenya versus Tanzania. *World Politics*, **56** (03), 328–362.

— and Gugerty, M. K. (2005). Ethnic diversity, social sanctions, and public goods in Kenya. *Journal of Public Economics*, **89** (11), 2325–2368.

Montalvo, J. and Reynal-Querol, M. (2005). Ethnic diversity and economic development. *Journal of Development economics*, **76** (2), 293–323.

Mwakikagile, G. (2010). *The Gambia and Its People: Ethnic Identities and Cultural Integration in Africa*. Dar es Salaam, Tanzania: New Africa Press.

Papke, L. E. and Wooldridge, J. M. (1996). Econometric methods for fractional response variables with an application to 401(k) plan participation rates. *Journal of Applied Econometrics*, **11** (6), 619–32.

Perfect, D. (2008). Politics and society in The Gambia since independence. *History Compass*, **6** (2), 426–438.

Quinn, C. A. (1972). *Mandingo Kingdoms of the Senegambia: Traditionalism, Islam, and European Expansion*. Northwestern University Press.

Saine, A. (2008). The Gambia's elected autocrat poverty, peripherality, and political instability, 1994–2006: A political economy assessment. *Armed Forces & Society*, **34** (3), 450–473.

Sallah, T. M. (1990). Economics and politics in The Gambia. *The Journal of Modern African Studies*, **28** (4), 621–648.

Swindell, K. (1978). Family farms and migrant labour: The strange farmers of The Gambia. *Canadian Journal of African Studies / Revue Canadienne des Études Africaines*, **12** (1), 3–17.

Thomson, S. (2011). Revisiting "Mandingization" in coastal Gambia and Casamance (Senegal): Four approaches to ethnic change. *African Studies Review*, **54** (2), 95–121.

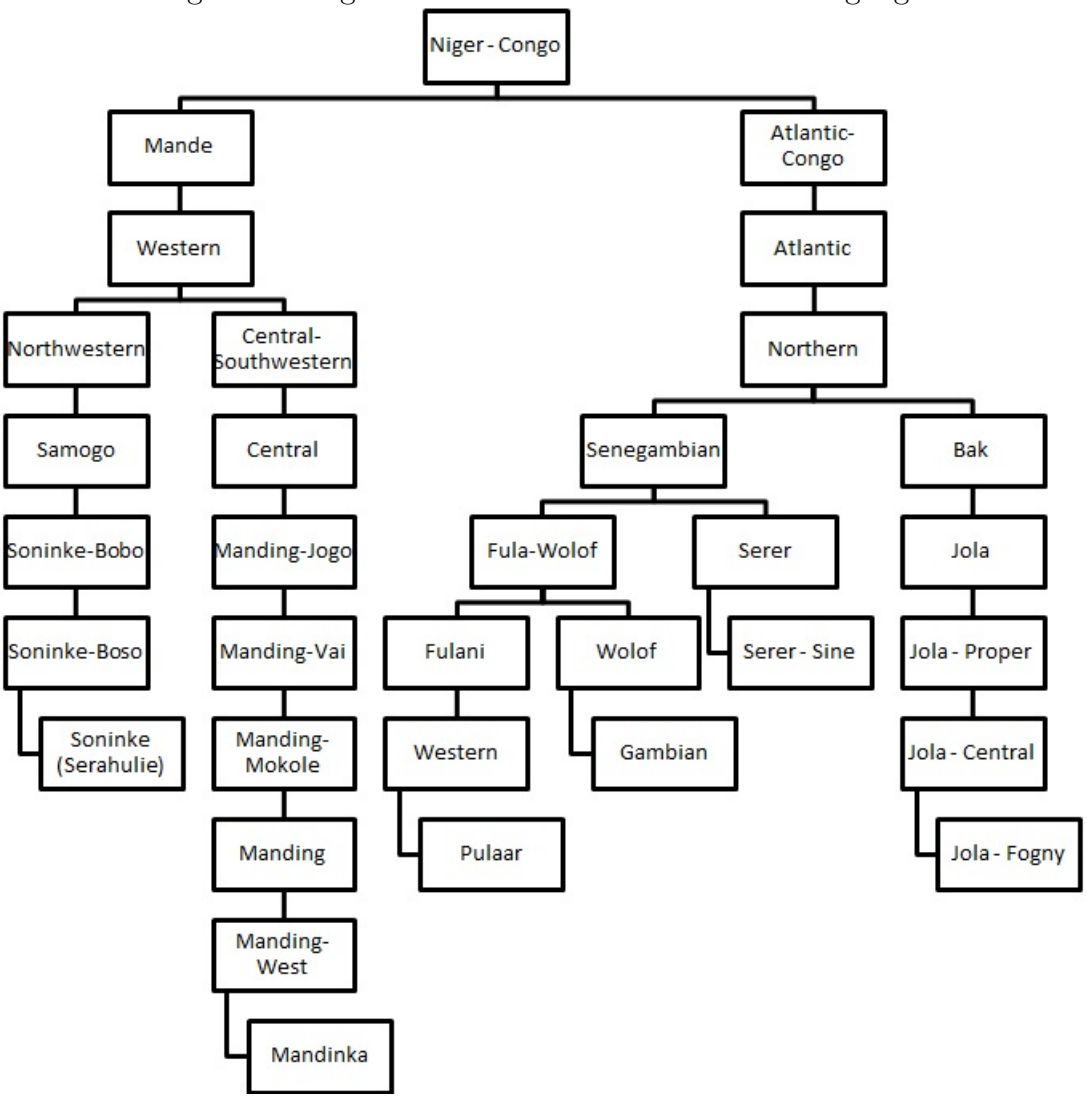
— (2012). Developing a multiethnic ethos: How colonial legacies, national policies, and local histories converged in a Gambian village charter. *Studies in Ethnicity and Nationalism*, **12** (2), 286–308.

World Bank (2013). World Development Indicators [Data file]. Retrieved from http://data.worldbank.org/data-catalog/world-development-indicators.

Wright, D. R. (1999). "What do you mean there were no tribes in Africa?": Thoughts on boundaries –and related matters– in precolonial Africa. *History in Africa*, **26**, 409–426.

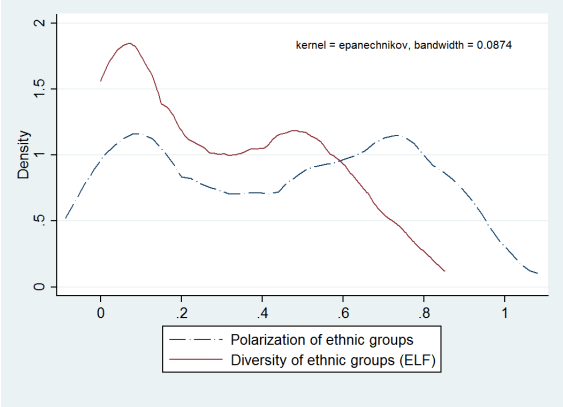
— (2004). *The World and a very small place in Africa: a history of globalization in Niumi, the Gambia*. ME Sharpe.

Figure 1: Linguistic tree of the main Gambian languages



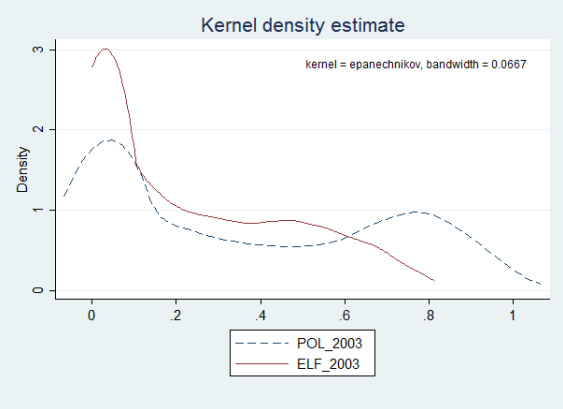
Source: Ethnologue (Lewis *et al.*, 2014)

Figure 2: Ethnic diversity and polarization in our data



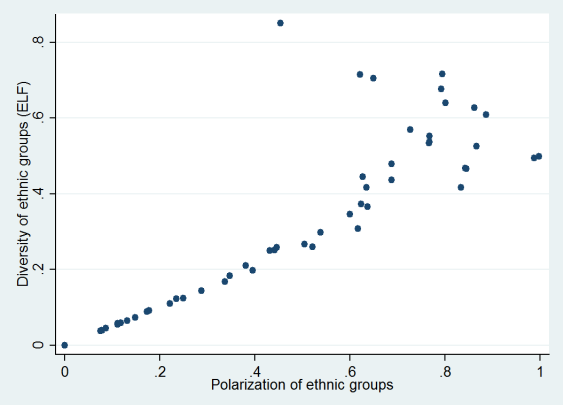
Note: Kernel estimation of the distribution of ELF and POL in the 59 Gambian villages of our sample. Data described in Table 2

Figure 3: Ethnic diversity and polarization at national level



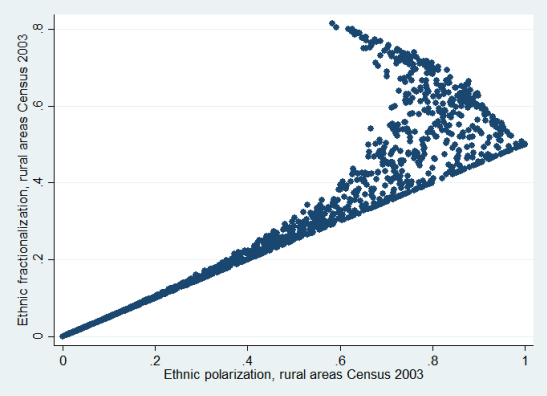
Note: Kernel estimation of the distribution of ELF_{2003} and POL_{2003} in all rural Gambian villages. Data from National Census 2003 (1811 villages).

Figure 4: Ethnic diversity and polarization in our data



Note: Ethnic polarization (POL) versus fractionalization (ELF) in 59 Gambian villages.

Figure 5: Ethnic diversity and polarization at national level



Note: Ethnic polarization (POL_{2003}) versus fractionalization (ELF_{2003}) in all rural Gambian villages. Data from National Census 2003 (1811 villages)

Table 1: ETHNICS GROUPS IN RURAL GAMBIA

Ethnic Group	Network Sample	Census 2003 All country	Census 1993 Sample	Census 2003 Sample
Mandinka	50%	33%	56%	52%
Fula	23%	25%	19%	21%
Wollof	10%	14%	8%	9%
Jola	8%	8%	6%	7%
Serer	5%	2%	5%	6%
Serahule	2%	11%	1%	2%
Others (including non-Gambians)	3%	7%	4%	2%

Note: Share in the total rural population of different ethnic groups. First column corresponds to our sample (59 villages). Column 2 is based on data from the Gambia National Census 2003 for all rural villages in the country (1,811 villages). Column 3 is based on data from the Gambia National Census 1993 for the villages in our sample (55 villages) and column 4 the same but using the 2003 census (59 villages).

Table 2: DATA DESCRIPTION

Variable	Mean	Std. Dev.	Min	Max	Source
VILLAGE LEVEL DATA					
Approximate population	587	247	202	1,402	Own
Population density (persons/km^2)	6,911	4,152	652	20,310	Own
Gini (from self-declared income)	0.34	0.11	0.13	0.66	Own
Diversity of economic activity (Herfindahl)	0.41	0.18	0	0.81	Own
Diversity in educational level (Herfindahl)	0.71	0.12	0.47	0.95	Own
Poverty index	0.650	0.094	0.354	0.921	Census 2003
Grass huts (village %)	38	29	0	94	Own
Migrants (%, more than 20 years old)	0.311	0.210	0	0.902	Census 2003
Ethnic Fractionalization (ELF)	0.29	0.24	0	0.85	Own
Ethnic Fractionalization 2003 (ELF_{2003})	0.324	0.263	0.004	0.982	Census 2003
Ethnic Fractionalization 1993 (ELF_{1993})	0.285	0.281	0.005	0.981	Census 1993
Polarization of ethnic groups (POL)	0.44	0.31	0	0.99	Own
HOUSEHOLD LEVEL DATA					
Household Size	12.67	11.40	1	400	
Age of household head	51.70	15.54	15	100	
Female Household head (%)	6	23			
Polygamous (%)	47	50			
Monogamous (%)	47	50			
Non-Muslim (%)	0.8	9			
Ethnic minority (< 30%)	0.17	0.38	0	1	
Ethnic minority (between 30% and 50%)	0.03	0.17	0	1	
Land per worker (hectares)	2.27	7.40	0	133	
Income per capita (GMD)	3,514	4,735	43	125,000	
Agricultural income (% of total)	12	24	0	100	

Note: Village-level data: 59 observations for each variable (except for census 1993 data, with 55 observations). "Own" implies data come from our database. Household-level data: 2,689 observations for each variable. All household-level data come from our data.

Table 3: NETWORKS DESCRIPTION

Network		HOUSEHOLD LEVEL		VILLAGE LEVEL			
		In-degree centrality	Out-degree centrality	Links	Density ($D(m)$)	Clustering ($Cl(m)$)	Compactness ($Cmp(m)$)
LAND	Mean	0.010	0.010	29.2	0.032	0.009	0.210
	s.d.	(0.019)	(0.035)	(16.6)	(0.030)	(0.028)	(0.214)
LABOR	Mean	0.014	0.014	29.9	0.030	0.046	0.253
	s.d.	(0.036)	(0.024)	(20.8)	(0.026)	(0.065)	(0.232)
INPUTS	Mean	0.021	0.021	39.8	0.048	0.067	0.325
	s.d.	(0.034)	(0.038)	(22.5)	(0.044)	(0.072)	(0.278)
CREDIT	Mean	0.009	0.009	23.0	0.022	0.034	0.159
	s.d.	(0.017)	(0.027)	(17.9)	(0.021)	(0.053)	(0.184)
MARRIAGE	Mean	0.018	0.018	49.9	0.046	0.059	0.285
	s.d.	(0.092)	(0.092)	(40.9)	(0.039)	(0.073)	(0.253)
KINSHIP	Mean	0.098	0.098	147.7	0.144	0.279	0.638
	s.d.	(0.091)	(0.091)	(76.5)	(0.105)	(0.134)	(0.257)

Note: 2,689 observations for household-level information. 59 observations for village-level information. 's.d.' refers to the standard deviation.

Table 4: NETWORK CHARACTERISTICS: VILLAGE-LEVEL POOLED DATA

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	DENSITY ($D(m)$)			CLUSTERING ($Cl(m)$)			COMPACTNESS ($Cmp(m)$)		
	All	Economic	Family	All	Economic	Family	All	Economic	Family
WITHOUT CONTROL VARIABLES									
ELF	-0.001	0.007	-0.015	0.039	0.054***	0.008	-0.006	0.113	-0.246*
	(0.029)	(0.018)	(0.052)	(0.027)	(0.017)	(0.068)	(0.094)	(0.105)	(0.134)
Within R^2	0.000	0.001	0.001	0.003	0.025	0.000	0.000	0.007	0.025
Observations	354	236	118	354	236	118	354	236	118
CONTROL VARIABLES INCLUDED									
ELF	-0.006	-0.000	-0.016	0.008	0.002	0.020	-0.010	0.049	-0.114
	(0.020)	(0.014)	(0.041)	(0.025)	(0.024)	(0.070)	(0.125)	(0.167)	(0.132)
log(population)	-0.052***	-0.004	-0.093***	0.002	0.037*	-0.030	-0.154***	-0.000	-0.173***
	(0.010)	(0.007)	(0.019)	(0.014)	(0.020)	(0.039)	(0.044)	(0.076)	(0.062)
log (household size)	0.040**	0.001	0.071*	0.036	0.020	0.038	0.214*	0.114	0.172
	(0.018)	(0.013)	(0.036)	(0.022)	(0.026)	(0.072)	(0.107)	(0.150)	(0.106)
log (income per capita)	-0.002	-0.000	-0.000	-0.001	0.009**	-0.016	0.031	0.038	0.029*
	(0.003)	(0.002)	(0.007)	(0.004)	(0.004)	(0.013)	(0.021)	(0.030)	(0.017)
Gini	0.052*	0.021	0.054	-0.053	-0.024	-0.172*	0.282*	0.253	0.096
	(0.030)	(0.016)	(0.058)	(0.033)	(0.035)	(0.094)	(0.153)	(0.177)	(0.161)
Poverty Index	0.150***	0.022	0.275***	0.015	0.044	-0.109	0.105	-0.288	0.098
	(0.053)	(0.043)	(0.097)	(0.059)	(0.097)	(0.156)	(0.241)	(0.378)	(0.305)
% grass huts	-0.002	0.022	-0.051	-0.026	-0.042	-0.023	0.171	0.463**	-0.328**
	(0.022)	(0.021)	(0.042)	(0.031)	(0.039)	(0.087)	(0.129)	(0.181)	(0.138)
% migrants	0.016	0.035***	-0.023	0.052***	0.107***	-0.062	-0.022	0.122	-0.294***
	(0.017)	(0.010)	(0.029)	(0.019)	(0.018)	(0.048)	(0.097)	(0.112)	(0.105)
% relatives		0.178***			0.121			0.960**	
		(0.038)			(0.085)			(0.470)	
Within R^2	0.634	0.578	0.672	0.644	0.340	0.668	0.504	0.460	0.687
Observations	354	236	118	354	236	118	354	236	118

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors are clustered at the village level.
OLS estimation. All the regressions include ward- and network-level fixed effects and the lower panel includes the following additional control variables: main village activity, predominant ethnic group, % external links in each network, semi-urban areas, population density, presence of Marabout, household size diversity, education level diversity, % of Alkalo's relatives, % female household heads, and average land per worker.

Table 5: NETWORK CHARACTERISTICS: VILLAGE-LEVEL BY NETWORK

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	DENSITY ($D(m)$)				CLUSTERING ($Cl(m)$)				COMPACTNESS ($Cmp(m)$)			
	LAND	LABOR	INPUT	CREDIT	LAND	LABOR	INPUT	CREDIT	LAND	LABOR	INPUT	CREDIT
WITHOUT CONTROL VARIABLES												
ELF	0.030	0.000	-0.012	0.008	0.014	0.077*	0.043	0.083*	0.312*	0.081	0.035	0.025
	(0.025)	(0.017)	(0.028)	(0.016)	(0.016)	(0.039)	(0.038)	(0.039)	(0.148)	(0.175)	(0.177)	(0.099)
Within R^2	0.045	0.000	0.002	0.006	0.008	0.051	0.013	0.083	0.081	0.004	0.001	0.001
Observations	59	59	59	59	59	59	59	59	59	59	59	59
CONTROL VARIABLES INCLUDED												
ELF	0.038*	-0.015	-0.022	0.026	0.045	0.083	0.017	0.080	0.481**	0.069	-0.086	0.194
	(0.021)	(0.027)	(0.030)	(0.018)	(0.045)	(0.047)	(0.063)	(0.059)	(0.188)	(0.303)	(0.222)	(0.163)
log(population)	-0.018	0.004	-0.017	-0.007	-0.046*	0.023	0.008	0.024	-0.219	0.033	-0.214	-0.017
	(0.012)	(0.010)	(0.014)	(0.013)	(0.022)	(0.027)	(0.042)	(0.046)	(0.152)	(0.139)	(0.130)	(0.132)
log (household size)	0.005	0.035*	-0.044	0.034	0.003	0.037	-0.038	0.038	0.252	0.537*	0.062	0.212
	(0.026)	(0.019)	(0.032)	(0.022)	(0.032)	(0.073)	(0.060)	(0.058)	(0.192)	(0.265)	(0.323)	(0.252)
Gini	0.063	0.004	0.040	0.047**	0.026	-0.152*	0.033	0.074	0.522	0.083	0.748**	0.377*
	(0.039)	(0.022)	(0.058)	(0.019)	(0.034)	(0.080)	(0.060)	(0.082)	(0.423)	(0.212)	(0.340)	(0.204)
% migrants	0.001	0.009	0.062***	-0.011	0.024	0.031	0.198***	0.002	-0.195	-0.180	0.239	-0.148
	(0.015)	(0.025)	(0.019)	(0.010)	(0.026)	(0.063)	(0.064)	(0.110)	(0.164)	(0.237)	(0.188)	(0.130)
% relatives	0.141*	0.110	0.340***	0.106	-0.064	0.134	0.383	0.045	0.154	0.792	0.820	1.043
	(0.076)	(0.069)	(0.101)	(0.078)	(0.107)	(0.213)	(0.261)	(0.219)	(0.579)	(1.091)	(0.969)	(1.048)
Within R^2	0.706	0.760	0.818	0.682	0.475	0.517	0.666	0.397	0.482	0.572	0.659	0.485
Observations	59	59	59	59	59	59	59	59	59	59	59	59

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors, clustered at the village level, are not shown.
OLS estimation. All the regressions include ward-level fixed effects and the lower panel includes the following additional control variables: Poverty Index, % grass huts , log (income per capita), predominant ethnic group, % external links in each network, semi-urban, presence of Marabout, percentage female household heads, % of Alkalo's relatives, and average land per worker.

Table 6: NETWORK CHARACTERISTICS: POLARIZATION AND ELF_{1993}

		(1) All networks	(2) Economic networks	(3) LAND	(4) LABOR	(5) INPUT	(6) CREDIT
POLARIZATION: OLS ESTIMATION							
$D(m)$	POL	-0.007	-0.003	0.025	-0.016	-0.016	0.017
		(0.014)	(0.012)	(0.019)	(0.026)	(0.029)	(0.014)
	R^2	0.660	0.587	0.764	0.782	0.793	0.682
$Cl(m)$	POL	0.008	0.003	0.025	0.065	-0.001	0.044
		(0.020)	(0.018)	(0.025)	(0.056)	(0.055)	(0.042)
	R^2	0.654	0.376	0.553	0.612	0.688	0.462
$Cmp(m)$	POL	0.011	0.014	0.303*	0.009	-0.030	0.171
		(0.080)	(0.102)	(0.177)	(0.237)	(0.204)	(0.148)
	R^2	0.558	0.494	0.584	0.641	0.699	0.545
	Observations	354	236	59	59	59	59
ELF_{1993}: OLS ESTIMATION							
$D(m)$	ELF_{1993}	0.026***	0.031***	0.068***	0.010	0.043	0.030
		(0.011)	(0.015)	(0.021)	(0.026)	(0.040)	(0.028)
	R^2	0.675	0.621	0.817	0.796	0.817	0.852
$Cl(m)$	ELF_{1993}	0.038	0.023	0.017	0.009	0.097	0.063
		(0.023)	(0.022)	(0.045)	(0.062)	(0.077)	(0.072)
	R^2	0.654	0.394	0.539	0.586	0.643	0.476
$Cmp(m)$	ELF_{1993}	0.227**	0.268**	0.595**	0.157	0.363	0.059
		(0.105)	(0.133)	(0.259)	(0.342)	(0.298)	(0.291)
	R^2	0.573	0.524	0.595	0.648	0.722	0.539
	Observations	330	220	55	55	55	55
ELF: INSTRUMENTAL VARIABLE ESTIMATION							
$D(m)$	ELF			0.099***	0.014	0.059	0.043
				(0.024)	(0.027)	(0.042)	(0.027)
$Cl(m)$	ELF			0.042	0.005	0.106	0.091
				(0.049)	(0.068)	(0.077)	(0.066)
$Cmp(m)$	ELF			0.865***	0.230	0.507	0.084
				(0.282)	(0.354)	(0.316)	(0.280)
	Observations			54	54	54	54
	Kleibergen-Paap F statistic			18.10	23.60	21.96	20.74

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Robust standard errors.
All the regressions include ward-level fixed effects and the same control variables as in Tables 4 and 5. The pooled network regressions (columns 1 and 2) also include network fixed effects. The instrumental variable in the lower panel is ELF_{1993}.

Table 7: ETHNIC MINORITY AND HOUSEHOLD'S DEGREE CENTRALITY

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
X_{iv}	LAND		LABOR		INPUT		CREDIT	
	Lender	Borrower	Lender	Borrower	Lender	Borrower	Lender	Borrower
Ethnic minority	0.001	0.005**	-0.000	0.001	0.003	0.003	0.004**	0.000
	(0.004)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)
Relatives (%)	0.035**	-0.007	0.024***	0.066***	0.091***	0.050***	0.075***	0.026***
	(0.016)	(0.008)	(0.009)	(0.018)	(0.020)	(0.014)	(0.015)	(0.008)
log(Household size)	0.001	0.004***	0.001	0.005**	0.004**	0.002	0.006***	0.001
	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)
log(Age)	-0.000	-0.001	-0.006***	0.004	-0.004	-0.008***	-0.004	0.000
	(0.003)	(0.001)	(0.002)	(0.003)	(0.004)	(0.003)	(0.002)	(0.001)
Education	0.003	-0.003**	0.001	-0.002	-0.001	-0.004**	-0.002	0.001
	(0.003)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
log(Income per capita)	0.000	-0.000	-0.000	0.000	0.000	0.000	0.001*	-0.000*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Cashcrop producer	0.001	0.002**	0.002	0.003**	0.004**	0.002	0.002	-0.000
	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.001)
Remitances	-0.001	-0.001	-0.002	-0.003**	-0.002	-0.001	-0.002	-0.000
	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)
Emigrants	0.001	0.004***	0.000	-0.001	0.004*	-0.003**	-0.001	-0.002*
	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)
Observations	2,689	2,689	2,689	2,689	2,689	2,689	2,689	2,689
Within R^2	0.167	0.092	0.062	0.099	0.138	0.061	0.129	0.064

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors, clustered at village level, in parentheses.
OLS estimation for the model presented in Equation 2. Village fixed-effect always included.
Other variables included in the regression but not reported due to lack of space are: ethnic group and main economic activity of the household head, percentage of active workers, compound head, agriculture as percentage of total income, access to newspapers and TV news, use of kamanyango system, proxy responddent, female household head, polygamous household, non-Muslim, external links to the village and traditional roles of the household head in the village. The complete results are reported in Jaimovich (2011).

Table 8: ETHNIC MINORITY AND HOUSEHOLD'S DEGREE CENTRALITY: ROBUSTNESS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	LAND		LABOR		INPUT		CREDIT	
	Lender	Borrower	Lender	Borrower	Lender	Borrower	Lender	Borrower
OLS ESTIMATES USING RELATIVE AND ABSOLUTE ETHNIC MINORITY								
Ethnic minority (≤ 30%)	0.002	0.004**	0.001	-0.000	0.004	0.003	0.004**	0.000
	(0.004)	(0.002)	(0.001)	(0.002)	(0.002)	(0.003)	(0.002)	(0.001)
Ethnic minority (31% - 50%)	-0.009**	0.011	-0.005	0.009	-0.002	0.004	0.004	0.001
	(0.004)	(0.008)	(0.003)	(0.007)	(0.005)	(0.003)	(0.003)	(0.002)
Relatives (%)	0.037**	-0.009	0.025***	0.065***	0.092***	0.050***	0.075***	0.026***
	(0.016)	(0.008)	(0.009)	(0.020)	(0.020)	(0.014)	(0.015)	(0.007)
Within R^2	0.168	0.095	0.063	0.100	0.138	0.061	0.129	0.064
OLS ESTIMATES USING PROPORTION OF ETHNICITY IN VILLAGE'S POPULATION								
% ethnic group	-0.008	-0.002	-0.001	-0.002	-0.007*	-0.006	-0.007**	-0.000
	(0.006)	(0.002)	(0.002)	(0.003)	(0.004)	(0.004)	(0.003)	(0.001)
Relatives (%)	0.040**	-0.011	0.026***	0.066***	0.094***	0.050***	0.075***	0.027***
	(0.016)	(0.008)	(0.009)	(0.020)	(0.020)	(0.014)	(0.014)	(0.008)
Within R^2	0.167	0.087	0.062	0.099	0.138	0.061	0.128	0.064
Observations	2,689	2,689	2,689	2,689	2,689	2,689	2,689	2,689

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors, clustered at village level, in parentheses.
OLS estimation for the model presented in Equation 2. Village fixed-effect always included.
The model includes dummies for predominant ethnic group and main economic activity of the household head (not reported). Other variables are included in the regression, but not reported due to lack of space: percentage of active workers, compound head, agriculture as percentage of total income, access to newspapers and TV news, use of kamanyango system, self-respondent interviewed, female household head, polygamous household, non-Muslim, number of emigrants, external links to the village and traditional roles of the household head in the village.

Table 9: DYADIC REGRESSION: ETHNICITY AND OTHER DYAD CHARACTERISTICS

	(1) ALL NETWORKS	(2) LAND	(3) LABOR	(4) INPUTS	(5) CREDIT
Same ethnic group	0.005*	-0.005***	0.007***	0.006***	-0.000
	(0.003)	(0.002)	(0.002)	(0.002)	(0.001)
Direct kinship relation	0.160***	0.018***	0.076***	0.091***	0.048***
	(0.006)	(0.003)	(0.005)	(0.005)	(0.003)
Kinship through marriage	0.058***	0.004	0.038***	0.024***	0.011**
	(0.011)	(0.005)	(0.008)	(0.007)	(0.005)
Sum household size	0.016***	0.004***	0.005***	0.005***	0.007***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Sum income *per capita*	0.002***	0.001**	0.001**	0.001	0.000
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
Difference income *per capita*	-0.001	-0.001**	-0.001**	-0.001	0.000
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
Difference land *per capita*	0.001***	0.002***	0.000	0.000	-0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Sum education	-0.001	-0.001	0.000	-0.005***	-0.001
	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)
Alkalo	0.087***	0.066***	0.021***	0.030***	0.037***
	(0.011)	(0.008)	(0.007)	(0.007)	(0.007)
Observations	70,007	70,007	70,007	70,007	70,007
Within R^2	0.041	0.013	0.024	0.027	0.016

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Two-way (i and j) clustered standard errors in parentheses. OLS estimates for the model presented in equation (3). Household fixed effects always included, as well as other sums and differences of other variables: percentage of active workers, compound head, cash crop seller, agriculture as percentage of total income, access to newspapers and TV news, use of kamanyango system, self-respondent interviewed, female household head, polygamous household, non-Muslin, number of emigrants, external links to the village, and traditional roles of the household head in the village.

Table 10: DYADIC REGRESSION: WITHOUT KINSHIP

	(1) ALL NETWORKS	(2) LAND	(3) LABOR	(4) INPUTS	(5) CREDIT
Same ethnic group	0.024***	-0.001	0.014***	0.006***	0.004***
	(0.003)	(0.002)	(0.002)	(0.002)	(0.001)
Observations	70,007	70,007	70,007	70,007	70,007
Within R^2	0.016	0.014	0.006	0.027	0.007

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Two-way (i and j) clustered standard errors in parentheses. Logit estimates for the model presented in equation (3). OLS estimates for the model presented in equation (3). Household fixed effects and the same control variables as in Table 9 are always included.

A Appendix A: Migration and ethnic diversity in the National Census 2003

The main objective of this appendix section is to analyze the data from the Gambian National Census 2003 in order to shed light about the determinants of the ethnic diversity in rural areas. In particular, we are interested on the effects of recent migration as a determinant of the village-level ethnic fractionalization index, ELF_{2003}, described in Table 2. Only the 435 rural villages which have less than 1,000 and more than 300 inhabitants are considered, since the sample used in the main analyses of the paper is representative of these villages (which are around 30% of the total rural Gambian villages).

From the census is possible to identify the number of persons who have migrated to the village with the question "*Were you born in this village?*". If the answer is negative, then it is possible to identify if the respondent was born in a rural area, an urban area, or outside of the country.[24] On average, 12% of the village inhabitants have immigrated from other rural areas, while 3% come from urban areas. Less than 1% come from another country, mainly Senegal.

The first column of Table A.1 shows that the variables the percentage of migrants according to the place of origin only explain a small fraction of the within-district variation of ELF_{2003} (R^2=0.057). The coefficient for migrants from urban and other rural areas are statistically significant, but only at the 10% and the estimated coefficients are relatively small: in both cases, a one standard deviation increase of the independent variable implies an increase of around 0.03 in the ELF index (its mean value in the sample is 0.3).

When control variables (village population, a dummy for the predominant ethnic group, and the percentage of people with access to improved water and percentage of people with some level of formal education), the second column of Table A.1 shows that the coefficients remain with similar magnitude but the percentage of urban migrants is not statistically significant anymore. Additionally, columns 3 to 8 show separate estimations for each of the six local geographical areas (LGAs) in rural Gambia. The results for *% migrants rural* suggest that most of the effect is driven by one region (LGA 4), as in the other LGAs the coefficient for this variable is mainly non-significant.

From these results it is possible to conclude that recent migration to the rural Gambian villages explain only a small fraction of the variation of ethnic diversity, suggesting that historical factors are likely to play an important role instead.

[24] In the analysis we exclude migrants declaring to have migrated in the last 4 months to the village in order to avoid the effect of temporary visitors, which are common in rural Gambia. The main results are robust to the inclusion of these migrants. This is also the case if only migrants with more than 20 years or 40 years old are considered.

Table A.1: ELF AND MIGRATION, CENSUS 2003

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
% migrants urban	0.420*	0.453	0.094	-0.629**	1.247*	9.627***	2.173**	4.742
	(0.237)	(0.294)	(0.238)	(0.166)	(0.522)	(1.714)	(0.422)	(2.923)
% migrants rural	0.288*	0.303*	0.402*	0.828***	0.059	0.120	-0.067	0.278
	(0.148)	(0.161)	(0.196)	(0.096)	(0.273)	(0.089)	(0.107)	(0.389)
% migrants international	1.282	1.029	-0.049	6.321**	0.715	0.057	5.507**	0.829
	(0.969)	(0.951)	(1.655)	(1.914)	(1.677)	(2.815)	(1.150)	(1.288)
CONTROLS	NO	YES	YES	YES	YES	YES	YES	YES
SAMPLE	ALL	ALL	LGA 3	LGA 4	LGA 5	LGA 6	LGA 7	LGA 8
District FE	YES	YES	YES	YES	YES	YES	YES	YES
Observations	435	435	71	45	108	59	74	78
Within R^2	0.057	0.076	0.120	0.583	0.118	0.357	0.206	0.252
Number of districts	34	34	9	6	6	5	4	4

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors are clustered at the district level.
The dependent variable is village-level ELF_{2003}, the ethnic fractionalization index calculated using data for all village inhabitants from the 2003 Gambia National Census.
OLS estimation using data from the 2003 Gambia National Census for rural villages with more than 300 and less than 1,000 inhabitants.
Control variables include village population, a dummy for the predominant ethnic group, and the percentage of people with access to improved water and percentage of people with some level of formal education.
Regressions 3 to 8 are separate estimations for each of the local geographical areas (LGAs) in rural Gambia.

B Appendix B: Robustness checks for main results

Table B.1: NETWORK CHARACTERISTICS: FRACTIONAL LOGIT REGRESSIONS

		(1) All networks	(2) Economic networks	(3) LAND	(4) LABOR	(5) INPUT	(6) CREDIT
Density	ELF	0.100	0.621	1.091**	0.274	0.148	1.649**
		(0.311)	(0.411)	(0.536)	(0.557)	(0.457)	(0.723)
	POL	0.060	0.297	0.629*	0.053	0.035	1.031**
		(0.212)	(0.280)	(0.368)	(0.386)	(0.299)	(0.473)
	ELF_{1993}	0.808***	1.292***	2.047***	1.039**	1.013**	1.158
		(0.224)	(0.302)	(0.455)	(0.463)	(0.457)	(0.736)
Clustering	ELF	0.205	0.816	.	2.925**	0.866	9.022***
		(0.322)	(0.760)	.	(1.348)	(0.858)	(3.395)
	POL	0.149	0.407	.	1.340*	0.336	3.883*
		(0.249)	(0.516)	.	(0.805)	(0.640)	(2.182)
	ELF_{1993}	0.512*	1.585***	.	2.878***	2.411**	1.247
		(0.288)	(0.522)	.	(1.088)	(0.945)	(2.033)
Compactness	ELF	0.225	1.202	3.005***	0.494	-0.616	2.093
		(0.666)	(0.775)	(0.741)	(1.126)	(0.910)	(1.276)
	POL	0.338	0.650	1.945***	0.294	-0.031	1.607*
		(0.420)	(0.504)	(0.637)	(0.782)	(0.591)	(0.872)
	ELF_{1993}	1.327***	1.856***	3.355***	1.336	0.984	0.576
		(0.502)	(0.571)	(0.807)	(1.121)	(0.925)	(1.332)

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors are clustered at the village level.
Fractional logit estimation. Each coefficient comes from a separate regression.
All the regressions include ward-level fixed effects and the same control variables as in Tables 4 and 5. The pooled network regressions (columns 1 and 2) also include network fixed effects.
The number of observations in each regression is the same as in the respective OLS estimations in the main text (59 villages for ELF and POL and 55 villages for ELF_{1993}).
The quasi-MLE estimation of the clustering equation for $LAND$ was not feasible (most values for the dependent variable are around zero).

Table B.2: VILLAGE CHARACTERISTICS FOR HIGH AND LOW VALUES OF ELF_{1993}

	LOW ELF_{1993}	HIGH ELF_{1993}	T-TEST DIFFERENCE p-value
Approximate population (log)	6.241	6.305	0.581
Average household size (log)	2.444	2.408	0.624
Average income per capita(log)	2.831	3.091	0.456
Land per worker (log)	0.458	0.291	0.729
Population density (persons/km^2)	8.752	8.541	0.249
Gini (from self-declared income)	0.320	0.355	0.276
Diversity of economic activity (Herfindahl)	0.103	0.102	0.969
Diversity in educational level (Herfindahl)	0.753	0.703	0.122
Poverty Index	0.661	0.654	0.777
Grass huts (village %)	0.344	0.443	0.215
Female Household head (%)	0.044	0.051	0.690
% external links	0.210	0.113	0.091
Migrants (%, more than 20 years old)	0.260	0.341	0.138
% relatives in the village	0.153	0.151	0.944

LOW ELF_{1993} show the mean value of village characteristics for the 28 villages with values below the median of ELF_{1993}, HIGH ELF_{1993} for 27 villages above the median.

Table B.3: OLS ESTIMATES OF ELF IN THE IV SAMPLE

	(1)	(2)	(3)	(4)
Dep. variable	LAND	LABOR	INPUT	CREDIT
Density	0.044	-0.028	-0.010	0.032
	(0.032)	(0.019)	(0.038)	(0.023)
R^2	0.706	0.791	0.817	0.702
Clustering	0.059	0.082	0.029	0.148**
	(0.050)	(0.060)	(0.069)	(0.054)
R^2	0.487	0.504	0.655	0.474
Compactness	0.508**	-0.013	0.067	0.217
	(0.223)	(0.291)	(0.244)	(0.209)
R^2	0.486	0.570	0.669	0.486
Observations	54	54	54	54

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Robust standard errors.
OLS estimates of the ELF coefficient in each of the network characteristics (dep. variable) for the sample in the instrumental variable estimates presented in Table 6.
All the regressions include ward-level fixed effects and the same control variables as in Tables 4 and 5.

Table B.4: ETHNIC MINORITY AND HOUSEHOLD'S DEGREE CENTRALITY

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
X_{iv}	LAND		LABOR		INPUT		CREDIT	
	Lender	Borrower	Lender	Borrower	Lender	Borrower	Lender	Borrower
Ethnic minority	-0.111	0.136***	-0.011	-0.004	0.041	0.058	0.191***	-0.011
	(0.071)	(0.042)	(0.043)	(0.041)	(0.041)	(0.037)	(0.074)	(0.032)
Relatives (%)	0.893***	-0.164	0.498***	1.111***	1.152***	0.687***	1.835***	0.645***
	(0.280)	(0.208)	(0.183)	(0.322)	(0.210)	(0.170)	(0.287)	(0.168)
Household size (log)	0.028	0.120***	0.031	0.126***	0.088***	0.018	0.196***	0.046
	(0.067)	(0.034)	(0.030)	(0.040)	(0.033)	(0.032)	(0.053)	(0.034)
Age (log)	-0.003	-0.048	-0.158***	0.147**	-0.062	-0.154***	-0.154*	0.006
	(0.085)	(0.044)	(0.049)	(0.065)	(0.062)	(0.047)	(0.087)	(0.054)
Education	0.046	-0.123***	-0.001	-0.074	-0.037	-0.096***	-0.075	0.005
	(0.065)	(0.041)	(0.030)	(0.055)	(0.038)	(0.035)	(0.048)	(0.045)
Income per capita (log)	0.006	-0.005	-0.003	0.009***	0.001	-0.002	0.013***	-0.011**
	-0.004	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.005)
Observations	2,689	2,689	2,689	2,689	2,689	2,689	2,689	2,689
R^2_{deviance}	0.404	0.301	0.288	0.330	0.386	0.371	0.338	0.296

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors, clustered at village level, in parentheses.
Fractional logit estimation for the model presented in Equation 2. Village fixed-effect always included.
The model includes dummies for predominant ethnic group and main economic activity of the household head (not reported). Other variables are included in the regression, but not reported due to lack of space: percentage of active workers, compound head, agriculture as percentage of total income, access to newspapers and TV news, use of kamanyango system, self-respondent interviewed, female household head, polygamous household, non-Muslim, number of emigrants, external links to the village and traditional roles of the household head in the village. The complete results are reported in Jaimovich (2011).

Table B.5: ETHNIC MINORITY AND HOUSEHOLD'S DEGREE CENTRALITY: HETEROGENEOUS EFFECTS

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	LAND		LABOR		INPUT		CREDIT	
	Lender	Borrower	Lender	Borrower	Lender	Borrower	Lender	Borrower
	OLS ESTIMATES							
Ethnic minority	0.037	0.015	-0.005	0.005	-0.018	0.002	0.061*	0.008
	(0.029)	(0.022)	(0.014)	(0.026)	(0.021)	(0.032)	(0.033)	(0.013)
Ethnic minority * relatives	-0.032	0.061	-0.029*	-0.056	-0.044	0.024	-0.020	-0.026
	(0.022)	(0.049)	(0.016)	(0.063)	(0.050)	(0.029)	(0.038)	(0.016)
Ethnic minority * household size	0.000	-0.002	0.002	-0.001	0.003	0.002	-0.002	0.000
	(0.003)	(0.002)	(0.002)	(0.003)	(0.004)	(0.004)	(0.004)	(0.002)
Ethnic minority * age	-0.008	0.000	0.001	-0.001	0.010	0.003	-0.007	-0.001
	(0.007)	(0.004)	(0.003)	(0.007)	(0.006)	(0.007)	(0.007)	(0.003)
Ethnic minority * education	0.009	-0.002	0.000	0.000	0.008**	0.001	-0.004	-0.002
	(0.008)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)
Ethnic minority * income per capita	-0.000	-0.000	-0.000	0.000	0.000	-0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
R^2	0.178	0.119	0.069	0.103	0.147	0.066	0.144	0.070
	OLS ESTIMATES USING PROPORTION OF ETHNICITY IN VILLAGE'S POPULATION							
% ethnic group	-0.031	-0.008	-0.004	0.017	0.036	-0.023	-0.074**	-0.017
	(0.041)	(0.018)	(0.020)	(0.039)	(0.034)	(0.034)	(0.035)	(0.018)
% ethnic group * relatives	-0.144***	0.033	0.018	0.014	-0.022	-0.062	0.065	0.011
	(0.046)	(0.026)	(0.029)	(0.070)	(0.083)	(0.053)	(0.047)	(0.021)
% ethnic group * household size	0.002	0.005	-0.001	0.005	0.002	0.001	0.005	0.002
	(0.005)	(0.003)	(0.003)	(0.005)	(0.005)	(0.005)	(0.005)	(0.002)
% ethnic group * age	0.005	-0.002	-0.001	-0.006	-0.017*	-0.004	0.008	0.002
	(0.011)	(0.004)	(0.005)	(0.010)	(0.009)	(0.009)	(0.007)	(0.004)
% ethnic group* education	-0.017	0.004	0.002	0.004	-0.010**	0.001	0.007*	0.001
	(0.010)	(0.003)	(0.004)	(0.005)	(0.004)	(0.005)	(0.004)	(0.003)
% ethnic group * income per capita	0.000	0.000	0.000	-0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)
R^2	0.184	0.100	0.069	0.103	0.148	0.069	0.143	0.071
Observations	2,689	2,689	2,689	2,689	2,689	2,689	2,689	2,689

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors, clustered at village level, in parentheses.
OLS estimation for the model presented in Equation 2. Village fixed effect always included.
The model includes dummies for predominant ethnic group and main economic activity of the household head (not reported). Other variables are included in the regression, but not reported due to lack of space: percentage of active workers, compound head, agriculture as percentage of total income, access to newspapers and TV news, use of kamanyango system, self-respondent interviewed, female household head, polygamous household, non-Muslim, number of emigrants, external links to the village and traditional roles of the household head in the village.

Table B.6: ETHNIC MINORITY AND HOUSEHOLD'S DEGREE CENTRALITY: RELATIVES IN THE VILLAGE NOT CONSIDERED

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	LAND		LABOR		INPUT		CREDIT	
	Lender	Borrower	Lender	Borrower	Lender	Borrower	Lender	Borrower
	OLS ESTIMATES							
Ethnic minority	-0.001	0.005**	-0.002	-0.002	-0.002	0.001	0.000	-0.001
	(0.004)	(0.002)	(0.001)	(0.002)	(0.003)	(0.002)	(0.002)	(0.001)
R^2	0.163	0.092	0.057	0.085	0.113	0.050	0.104	0.055
	OLS ESTIMATES USING RELATIVE AND ABSOLUTE ETHNIC MINORITY							
Ethnic minority (≤ 30%)	0.000	0.004**	-0.001	-0.004**	-0.002	0.001	-0.000	-0.002
	(0.005)	(0.002)	(0.001)	(0.002)	(0.003)	(0.003)	(0.002)	(0.001)
Ethnic minority (31% - 50%)	-0.009**	0.011	-0.006*	0.008	-0.003	0.003	0.002	0.000
	(0.004)	(0.008)	(0.003)	(0.007)	(0.005)	(0.003)	(0.003)	(0.002)
R^2	0.164	0.094	0.058	0.087	0.113	0.050	0.104	0.055
	OLS ESTIMATES USING PROPORTION OF ETHNICITY IN VILLAGE'S POPULATION							
% ethnic group	-0.004	-0.004	0.001	0.004	0.001	-0.001	-0.000	0.002
	(0.006)	(0.002)	(0.002)	(0.003)	(0.005)	(0.004)	(0.003)	(0.002)
R^2	0.163	0.086	0.056	0.085	0.112	0.050	0.104	0.055
Observations	2,689	2,689	2,689	2,689	2,689	2,689	2,689	2,689

Note: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors, clustered at village level, in parentheses.
OLS estimation for the model presented in Equation 2. Village fixed-effect always included.
The model includes dummies for predominant ethnic group and main economic activity of the household head (not reported). Other variables are included in the regression, but not reported due to lack of space: percentage of active workers, compound head, agriculture as percentage of total income, access to newspapers and TV news, use of kamanyango system, self-respondent interviewed, female household head, polygamous household, non-Muslim, number of emigrants, external links to the village and traditional roles of the household head in the village.