Quality of Social networks and educational investment decisions (Preliminary version)

Blanca Zuluaga
CES - University of Leuven

February 27, 2009

Abstract

All individuals belong to a social network with certain quality level. This paper analyses the role of the quality of the social network in the educational decision making process. We propose a measure for quality of netwok based on the schooling level and the labor position of the members of the net. our analysis compares individuals similar in at least two characteristics: socioeconomic level and intellectual ability. Although they belong to the same type of comunity (poor), they differ in the composition of their social network. The higher the quality of the network, the higher the probability of investing in education. Hence, socially disadvantaged and equally inteligent individuals may end up in different schooling equilibrium levels.

1 Introduction[1]

It is well known that certain factors like ability and the level of family income play an important role in determining the amount of human capital investment that an individual is willing to undertake. Although public education is free in many countries at basic and medium level of schooling - there is no monthly fee- and relatively cheap at the superior level, there are other costs like transport, food and clothes, among others, that poor families cannot afford. Besides, credit markets are incomplete, excluding most of the low income potential applicants.

As Roemer (1993) points out, the opportunity cost of being at school is lower for the rich than for the poor, making the investment in education less costly for the wealthier. Besides, a poor person may (perhaps correctly) believe that compared to an equally educated but rich person, there would not be many

[1]I would like to thank my supervisor Erik Schokkaert for his invaluable guidance. I am also grateful to Rocio Vera and Lorena Collazos from Universidad Icesi for their contribution to the discussions and the field work.

jobs attainable for him. This belief may discourage him from carrying out high levels of educational investment, for the instrumental value of education is not actually perceived.[2]

In this paper we want to explore the existence of an additional factor influencing the decisions on investment in education: the social networks[3] that individuals belong to. In particular, we state that the "quality" of the network influences the expected returns to education, encouraging or discouraging individuals investment in additional years of schooling. The "quality" of the social networks refers to the schooling level and links with the labor market of the network's members (e.g. family members, neighbors, colleagues, chiefs, classmates, teachers, among others).

It is possible to find individuals belonging to the same community or neighborhood, who share certain attributes like family income and ability, ending up at different equilibrium levels of schooling, expected future income and expected social mobility.[4] The analysis of social networks may offer us good ideas to explain this phenomenon and to explore why policies of educational expansion favor only a small portion of low income individuals.

Social networks are important in shaping the aspirations and expectations of individuals. These, in turn, affect the behavior and decisions of people in crucial aspects such as investment in human capital. In general, educational achievements of members of poor communities generate a high positive multiplicative effect, this is, an effect beyond the private return of the educated individual.

The existing economic literature on social networks and their effect on problems resolution, on the decision making process and on the level of socioeconomic achievements of the individuals, is very extensive. There also exists a vast liter-

[2] Education has both intrinsic and instrumental value. The first refers to the value of acquiring knowledge itself and the second refers to the positive influence of education in (among others) the capacity of individuals to get higher economic positions. If the individual experiences higher opportunity cost of schooling and lower chances to get a good job compared to an equally educated wealthier person, this affects his perception of the instrumental value of education.

[3] There is no a standard definition of social network in the literature. Jackson, (2005) defines it as the group of people "with whom we share information and favors on a regular basis"; According to Requena (2003), a social network is a set of social actors linked to each other through a number of relationships with properties like intensity of the relation, position of the actor, and accesibility of the actor with respect to the others.

[4] In the lenguage of Akerlof (1997), in poor neighborhoods some individuals follow the "status seeking behavior" while others follow "the conformity behavior". What does make the difference? It would be interesting to further analyse how the quality of the network influences the behavior of individuals and makes them status seekers or conformist. Although we will not work on this topic here, we may roughly state that, comparing individuals from the same poor neighborhood, those whose network's quality is (in average) higher follow a double strategy. They perform i) a status seeking behaviour w.r.t. the "low quality" members of their network and ii) a conformity behaviour w.r.t the high quality members of their network. An individual chooses x (some outcome) to maximize his utillity,

$$U = -d(\overline{x} - x) - h \mid \underline{x} - x \mid -ax^2 + bx + c$$

where $\overline{x}$ is the average outcome of low quality members and $\underline{x}$ is the average outcome of high quality members. The agent looses utility if $\overline{x} > x$ or $\underline{x} \neq x$. U is a function combining equations (1) and (3) of Akerlof (1997) p. 1008-1009.

ature on the formation of networks and their efficiency and stability conditions (See Jackson (2005) for a good review). This paper will not focus on how social networks arise nor on their nature, but on how they influence the educational investment decisions. Table 1 shows a few previous contributions to the discussion about the influence of social interactions on educational decisions. The order of the papers on table 1 reflects only the relevance of each paper for our own purposes.

Table 1: **Some previous work on the influence of social interactions on schooling investment**

Author	Channel	Main idea
Calvó-Armengol and Jackson (2005)	Correlation of human capital among generations.	The higher the S level of R's members, the higher i's expected α. Sensitivity of i's decisions w.r.t. R's composition determines perpetuation of inequality.
Streufert (2000)	Social isolation of low classes	Poor youths lose high income role models: they observe a distribution of Y truncated above. Hence, they underest. relation between S and Y.
Anderberg and Anderson (2007)	Social environment (SE)	If i's neighbors have 'Good' positions (high wages), i perceives higher α and invest more in S. SE is the only channel transmitting success among generations.
Moizeau et. al. (2004)	"Information effect" neighborhood	Children form their idea on α from experiences of older generations. Erroneous perception of α implies that indiv. from poor neighborhoods lack incentives to invest in S.
Yamauchi (2005)	Social learning and neighborhood effects	i learns about α by observing the Y level of their neighbors. The author employs a Bayesian model of learning where people decide on S subject to subjective uncertainty on α.
Durlauf (1992)	Incentives and aspirations	Parents´ election of neighb. determines characteristics of i's role models who influence aspirations & expectations of their children, and available funds to finance S in the community.

Y: income, S: schooling or years of schooling
R: social network, α: returns to education, i: individual

The studies shown in table 1 compare (theoretically and/or empirically) groups of population from different communities with different socioeconomic levels. In this paper we intend to compare groups inside the same type of

community, i.e. the poor. These individuals, in spite of belonging to the same neighborhood, partially differ in the composition of their social network. This type of focus applied to the decisions in education constitutes a contribution to the existing economic literature in social networks. In the second section of the paper we make explicit the way in which the quality of the social network affects the educational investment of individuals. The network's quality has an effect on the perception of the individual on the returns to education, which in turn influences his educational decisions. We propose a specific measure for the quality of social networks, whose information requirements are: i) "quality" of each member of the network, based on educational level and labor position, and ii) the weight of each member.

An important definition is the "key tie". This is a concept characterizing a member of the network who plays a decisive role in determining the overall quality of the social network. A key tie might be a strong or a weak tie; although these two concepts are commonly used in the literature, there is no consensus on their precise definition. We will not pretend here to be more accurate in defining the concepts, instead, we will adopt an ordering for strength of ties originally based on family (stronger), friends and acquaintances (weaker), and subsequently modified by factors like closeness, intimacy, economic support and admiration. These adjustment factors may lowered the weight of originally strong ties and, especially the two last factors, could make a weak time become a key tie.

We present a model that describes the individual decision making process. We claim that it is suitable to use a Logit model in order to check whether, as we expect, the variable "quality of the network" has a positive and significant influence in the decision of the individuals about to continue studying or not.

The third section corresponds to the empirical part. We work with two groups: treatment and control. Individuals in both groups are similar in their socioeconomic conditions and intellectual ability. Those in the first group have continued studying after secondary whereas those in the control group have not. We find out the schooling level and the labor position of each member of the individuals' network in order to estimate the quality of the network. We then presents the results of the Logit model to test the influence of the network quality on the decision of individuals to continue studying. Finally we present some conclusions and recommendations.

2 Educational investment and social networks

Let $N = \{1, \ldots, n\}$ be a set of agents living in the same neighborhood – poor-. In spite of sharing the same neighborhood, the social network R that each agent belongs to is not the same for all agents – although they inevitably do share part of their network. Let R_i be the social network of individual i. There is a quality level associated to each social network. R_i^q expresses that individual i belongs to a social network of quality q. Definition 1 (below) explains more precisely how to determine the quality of a network.

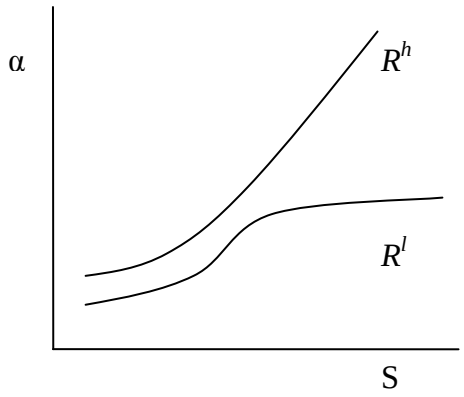

Figure 1: Expected Returns to Education

In our analysis, individuals are similar in at least two characteristics: they are intelligent (enough to carry out high educational levels), and they have a disadvantaged social and family background. They decide the amount of educational investment they want to undertake (S) – latter, we define S as a binary variable indicating if an individual invests or not in higher education. This decision is based on their perception of the returns to human capital investment (α). The expected returns to education are closely related to the nature of the social network that each individual i belongs to. Hence, we may express the returns as follows

$$\alpha_i = \alpha(R_i^q) \tag{1}$$

Individuals from low quality networks (R^l) will form an expectation of the returns to education that is lower than the expectation of individuals from high quality networks (R^h), and that will not always increase with years of education. Figure 1 illustrates this point.[5]

Why does the quality of the social network affect the expected returns to education? There are several channels of influence. Let us mention at least four: expected attainable jobs, expected future income, aspirations, and motivation. First, an individual belonging to low quality network may (perhaps correctly) believe that his chances to get a good job are lower than some one with better connections, which discourage him from investing in education. Second, individuals have an idea about the relationship between education and income, which is based on what they observe of the sample of individuals accessible to them. Hence, if an individual belongs to a low quality network R^l, his perception of the relation between education and income would be based on a sample that excludes high income observations at each schooling level.[6] Third, a poor person belonging to a low quality network may fail in detecting how education will

[5] The difference in the quality of the social network may also influence the individuals ´expected probability of failure in educational achievements.

[6] In the fashion of Streufert (2000), Let $F_s(y)$ be the cumulative income distribution func-

positively affect his welfare, which pulls down his educational aspirations. This aspiration trap (See Heifetz and Minelli (2006)) arises because they have few opportunities to experience how educational choices influences their "fundamental well-being". Finally, the quality of the social network has also an influence on the motivation level of individuals towards educational investment. Attending school requires an effort that only sufficiently motivated individuals are willing to exert. Achievements of high quality role models enhances individual's motivation and willingness to exert effort. Individuals feel encouraged by the possibility to catch up with their high quality role models.

2.1 Quality of networks

Let $M = \{1, \ldots, m\}$ be a set of role models. r^q_{ix} is the quality of i's role model r_x. $r_{ix} \in R_i$ and $R_i = \{r_{i1}, r_{i2} \ldots r_{im}\}$. The value associated to r^q reflects the educational level and link with the labor market of role model r_x, where $0 < r^q \leq 1$. The closer r^q is to 1, the higher the "quality" of the role model. Thus, the quality of the network will depend on the quality of the role models (ties or members of the network).[7]

Each r_{ix} may have a different weight (θ_x) in determining the quality of the social network, which will be denoted by R^q_i. The weight depends on the relative importance of the members in the network. There are difficulties in determining the relative importance of members though, since this is a subjective matter guided by emotional attachments, power relationships, among other factors. We will work out later on the determinants of the weight.

Definition 1 (Quality of a network) *Let each role model r_x to have a weight θ_x. The quality of a networks is denoted by R^q_i and is determined as follows*

$$R^q_i = \sum \theta_x r^q_{ix} \tag{2}$$

Where $\theta \in [0, 1]$, $\sum_{x=1}^{m} \theta_x = 1$ and $0 < R^q_i \leq 1$.

We will see that the size of the network (number of members) is also relevant in determining educational decisions, however, the size does not enter in the proposed network quality estimation.

Definition 2 (High quality network) *A social network is said to be high quality R^h_i if*

$$\sum \theta_x r^q_{ix} > \sigma \tag{3}$$

Thus, there is a value between 0 and 1 reflecting the network's quality. If this value surpasses σ, we say that individual i belongs to a high quality network.

---

tion – the probability that an individual investing s will get an income less than or equal to y -. An individual observes the "role models" or members of his social network R^q and learns about F_s. What happens to those individuals belonging to a low quality network R^l? They would learn about a cumulative income distribution that is truncated from above $F_s(\bar{y})$, since they do not observe any role model whose income surpasses a certain amount.

[7] We will use these three terms indistinctly through the paper.

The value of σ can be determined through a sample of individuals with information of their network's quality and educational level. Individuals with high schooling levels most likely belong to a network whose quality is above σ. We deal with its estimation in the empirical section. In general terms, from a continues approach, we would say that the higher the quality of the network, the higher the schooling level of the individual.

As we mentioned, individuals form their expectations of the returns to education α influenced by the social network they belong to. Those living in a poor neighborhood may differ in the quality of their social network, hence, in their perceptions about α. Let $i, j \in N$, with network's quality denoted by R_i^q and R_j^q respectively. Sharing the same neighborhood makes it likely for i and j to share some of the role models of their networks. However, there is at least one role model $r_{ix} \in R_i$ who does not belong to R_j. and makes the difference. We claim that the presence of at least one high quality member with high enough θ may positively influence the educational decision. We call this member a 'key tie' or 'key role model'. Let us r_{ik} denote the key tie. There might be more than one key tie.

Definition 3 (Key tie) *The key tie or key role model crucially helps in determining the quality of the social network. To understand "crucially", consider that there might be the case in which, for individual i,*

$$R_i^q < \sigma$$

whit $R_i = \{r_{i1}, r_{i2}, ...r_{im}/r_{ix} \neq r_{ik}\}$. However, considering the key tie,

$$R_i^q > \sigma$$

whit $R_i = \{r_{i1}, ..., r_{ik}, ..., r_{im}\}$.

r_{ik} has two characteristics i) he is a no family member 'high quality' tie, ii) his weight θ, initially lower than those of relatives and closest friends, ends up being as high or higher than originally stronger ties due to different adjustment factors (e.g. intimacy, admiration, emotional or economic support), that will be explained in section 2.3. r_{ik} is crucially helping in enhancing motivation, aspirations, and in general, non-cognitive skills of the individual.

But why could we fail in considering the key tie? Because, as suggested, the key tie may be an initially weak tie (acquaintances), with low θ if we were not adjust the weight considering other factors different to kinship. For individuals living in a poor neighborhood, strong ties (family, closest friends) are likely to be low-educated, so there might be some cases in which a weak tie turns up to be a key tie influencing individual decisions..

This is important in the context we are analyzing. Our individuals share the same poor neighborhood. Their strong ties are normally of similar characteristics. If poor and/or low-educated people only interact with their equally disadvantaged strong ties, the possibility of widening their life perspectives is lower. Granovetter (1983) points out that "the heavy concentration of social

energy in strong ties has the impact of fragmenting communities of the poor into encapsulated networks with poor connections between these units". The problem of these encapsulated social networks is that poor individuals lose the potential advantages of wider information received through their weak ties' own networks, which may be, according to the author, "one more reason why poverty is self-perpetuating".

In some cases, admiration for someone different to our strong ties may exert an important role in shaping our behavior.[8] In this sense, as important as the schooling level of the parents might be the presence of one or several high quality members in the network. Hence, it is not necessary that poor individuals observe many high quality role models (a complete reference group) in order to change their fate, a key tie may suffice.

2.2 The decision

Let $S = 1$ if individual i reaches high educational levels, $S = 0$ otherwise. We want to show that the probability that $S = 1$ increases when $i \in R^h$ and the probability that $S = 0$ increases when $i \in R^l$ (or $S(R^h, \alpha_i) = 1$ and $S(R^l, \alpha_i) = 0$). This is consistent with our claim that $\alpha_i(R^h, .) > \alpha_j(R^l, .)$, where $i, j \in N$. As explained above, a high quality network offers advantages to the individuals such as higher information on the true returns to education, enhancement of motivation to study and higher capacity to aspire, among others.

Figure 2 summarises the process through which the quality of the network influences the individual' decisions on education. From the center of the figure to the left, each individual i belongs to a social network R_i. This network is composed by high quality members (r^h_{ix}), low quality ties (r^l_{ix}) and key tie(s) (r_{ik}), whose weight is denoted by θ_x and θ_k. The weighted sum of the quality of the members determines R^q_i (definition 1). In turn, the quality of the network is a variable affecting the expected returns to education α of individual i.

What he observes of the members of his network (success, labor position, educational level and social mobility) and the motivation and capacity to aspire that his environment pass him down, contributes to the formation of the individual's perception of the return to education α. This return might be thought in a broad sense, this is, it does not only correspond to the expected monetary benefit of education (or any other instrumental value of education), but also to the intrinsic value or the value of knowledge itself. The higher the expected return to education, the higher the probability of deciding to invest in schooling.

There exist a threshold value of network quality above which individuals decide to invest in education. Thus, we may work out a Logit model in order to verify that S will be most probably equivalent to 1 for those individuals whose quality network is above σ (definition 2), and most likely equal to 0 for

[8] Weak ties refers to acquaintances and excludes media personalities. Members of the networks are those with whom there actually exist some type of real interaction. Media personalities might influence behaviour - for instance, a national tenis champion may encourage youngsters to enrroll to tenis schools - but we can hardly claim that they belong to the social network of all individuals who admire them.

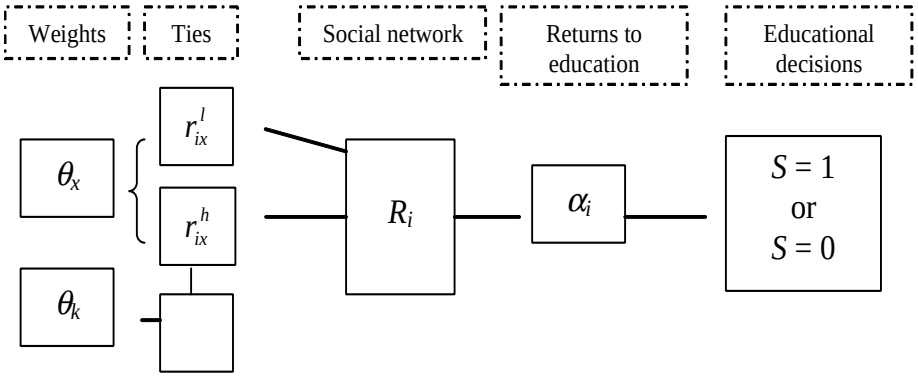

Figure 2: Summarising the influence of que quality of networks on educational decisions.

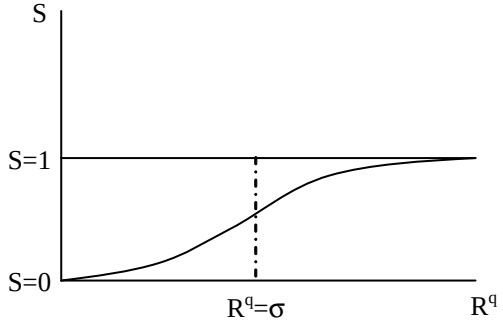

Figure 3: Quality of Network and educational investment decision

individuals belonging to networks with quality lower or equal to σ. Figure 3 illustrates this point.

The Logit model has the following form,

$$S_i = \beta_0 + \beta_1 R_i^q + \beta_2 \mathbf{X_i'} + \boldsymbol{\varepsilon_i} \tag{4}$$

Where $\mathbf{X}$ is a vector compiling other variables that might affect the schooling decision, such as number of siblings and size of the network (income level and ability are similar for all individuals). We expect β_1 to be significant and positive.

2.3 How θ is defined

Each individual i belongs to a network with m role models. A role model x has a weight θ_x, when determining the quality of the network R_i. We have that $\theta \in [0,1]$ and $\sum_{x=1}^{m} \theta_x = 1$. There are basically three categories of role models: family, friends and acquaintances. In principle we assume that θ is higher for family, followed by friends and finally acquaintances. However, there are certain factors like intimacy or closeness, emotional support, financial support, admiration, confidence, among others, that make θ differ among the members' subsets of family, friends and acquaintances. For instance, admiration may enhance the weight of apparently weak ties (acquaintances). An appropriate definition of θ_x might reveal cases in which parents background may loose importance in influencing individuals' decisions.

Suppose there is an individual i whose social network is composed by 5 relatives, 4 friends and 3 acquaintances. Let us θ_r, θ_f, and θ_a denote the weight of relatives, friends and acquaintances respectively, where initially we assume that $\theta_r > \theta_f > \theta_a$. Thus, the adjustment process of the weights consists of identifying members whose importance in the network may vary due to the factors mentioned above (support, intimacy, etc.). Table 2 shows an example of how the initial assigned weights to each member of the network (column 2) should be modified to consider the influence of those factors. Column 3 indicates that the weight of relative 3 and friends 2 and 4, should be adjusted (increased) because they give financial support to our individual. Column 4 indicates that friend 1's weight should be increased because our individual receives emotional support especially from him. Columns 5 and 6 indicates additional adjustments. Column 7 shows the number of adjustment factors that applies to each member. For them, the weight will be increased and, in order to comply that $\sum_{x=1}^{m} \theta_x = 1$, the weight of all the rest of ties should be lowered proportionally. We will call the factors from columns 3 to 6 'adjustment factors'. Finally, column 8 shows whether the initial θ has increased or decreased due to the relevant factors.

In summary, θ may be written as follows:

$$\theta_x = \theta(relationship_{(kin,friend,acquaintance)}, AES, AEmS, AI, AA) \qquad (5)$$

where variables are defined as in table 2. There are at least two questions to solve now: how to assign the initial weights and how to determine the impact on the initial θ of the adjustment factors in order to get the adjusted θ. Available literature sheds little or no light about this problem. However, to help ordering ties according to their strength, there are important contributions especially by Granovetter (1983) and Marsden and Campbell (1984). Granovetter suggests that "the strength of a tie is a (probably linear) combination of the amount of time, the emotional intensity, the intimacy (mutual confiding) and the reciprocal services which characterize the tie". Marsden and Campbell were inspired by this intuitive definition, going further in the effort to empirically test the best indicators for strength. They argue that a measure of closeness or emotional intensity of the relationship is the best indicator of the strength of ties, in

Table 2: **Adjustment of the weight**

Member	Initial θ	AES	AEmS	AI	AA	#AF	Adjusted θ
Relative 1	θ_{r1}					0	$\theta_{r1} \downarrow$
Relative 2	θ_{r2}					0	$\theta_{r2} \downarrow$
Relative 3	θ_{r3}	$\uparrow$				1	$\theta_{r3} \uparrow$
Relative 4	θ_{r4}			$\uparrow$		1	$\theta_{r4} \uparrow$
Relative 5	θ_{r5}					0	$\theta_{r5} \downarrow$
Friend 1	θ_{f2}		$\uparrow$		$\uparrow$	2	$\theta_{f2} \uparrow$
Friend 2	θ_{f2}	$\uparrow$				1	$\theta_{f2} \uparrow$
Friend 3	θ_{f3}					0	$\theta_{f3} \downarrow$
Friend 4	θ_{f4}	$\uparrow$				1	$\theta_{f4} \uparrow$
Acquaint. 1	θ_{a1}					0	$\theta_{a1} \downarrow$
Acquaint. 2	θ_{a2}				$\uparrow$	1	$\theta_{a2} \uparrow$
Acquaint. 3	θ_{a3}				$\uparrow$	1	$\theta_{a3} \uparrow$
$\Sigma\theta$	1						1

AES: Adjustment for economic support (besides parents)
AEmS:Adjustment for emotional support
AI: Adjustment for intimacy
AA: Adjustment for admiration
#AF: Number of adjustment factors

comparison with other suggested measures like "breadth of discussion topic" and mutual confiding. They emphasise that measures related to time spent in a relationship like frequency and duration of contact are not good indicators of tie strength, since they overestimate the strength of neighbors/coworkers and relatives respectively.

In addition, Marsden and Campbell analyse how accurate is to assume that relatives are strong ties and neighbors/coworkers are weak ties, concluding that, although the assumption is accurate, the results also show that "the combined ability of the predictors to account for strength is limited". Hence, they recommend to focus in closeness when determining strength of ties.

Considering these contributions and dealing with the lack of additional literature, in the next section we tell how, for the empirical application, we assign the initial θ and how we use the adjustment factors to get the adjusted θ.

3 Some empirics

3.1 The Survey

> "Given that identification based on observed observation is so tenuous, experimental and subjective data will have to play an important role in futures efforts to learn about social effects" (Manski 1993)

We apply a survey to individuals from a "treatment group" and a "control group". The first group (60) is composed by university students coming from poor neighborhoods. These students are admitted based on their good academic performance. The control group (35) is a sample of individuals similar in ability and socioeconomic background to those in the treatment group, but not enrolled (nor planning to enroll) in superior education programs. We made sure that the reason for individuals of the control group not to continue studying after secondary school was not lack of intellectual ability. We did it by checking their school degrees and results in the National test to access higher education (Icfes).

The survey help us to detect, for each individual, i) the quality of each of the members of his/her network, i.e. educational level and labour position, ii) the weight of each member, by identifying if individual receives emotional support from them, confidence, intimacy and admiration, plus some other indicators like frequency and type of contact and existence of financial support. With i) and ii) we are able to estimate the quality of the social network for all individuals in our sample, according to equation 2.

The social network is composed by parents, siblings, closest cousins, closest aunts and uncles, closest family in law, closest friends, closest professors, godparents, sentimental partners, closest neighbors, closest mates from associations or organizations, and relevant acquaintances. See the questionnaire at the appendix B.

3.2 The weights

We start by assigning uniform values of θ to each member of the network.

$$\theta_{uniform} = 1/m$$

Recall that m is the size of the network or number of ties.

Next, we apply a first adjustment factor (af_1) according to the type of relationship: relative (r), friend (f), acquaintance (a)

$$\theta_{initial} = \theta_{uniform} * af_1$$

with $\theta_r > \theta_f > \theta_a$.

After, we consider additional adjustment factors, in line with what we have found in the literature of strength of ties: economic (af_2) and emotional (af_3) support, confidence (af_4) and admiration (af_5). Previous work on the field does not tell us how to determine the impact of each adjustment factor on the initial weight. A more demanding field work than the one that we were able to do here would be required for this matter.

Our option was to assume that all adjustment factors have a uniform impact ($af_2 = af_3 = af_4 = af_5$) and pick up an arbitrary small factor to apply to $\theta_{initial}$. For instance, if there is a member of the network to which the four adjustment factors apply, we would obtain his weight as follows

$$\theta_{adjusted} = \theta_{initial} * af * 4$$

After the adjustment process we could eventually get a new ordering of the members' $\theta's$ in which an initially weaker tie (e.g. acquaintance) may ended up with a weight as important as the weight of a family member (or even more important), when the adjustment factors apply to this tie. In this way we can actually capture the key influence of initially weak ties on the social network.

In order to verify that the results shown in the next section are robust, we run the Logit model with three different calculations of the variable "quality of the network". The three calculations follow definition 1, but they differ in the weight (θ) applied to each member of the network (lets us call them R, $R2$ and $R3$). R was calculated using the weights just as it is described in this section and the results of the Logit model appear in the next section. $R2$ was calculated using uniform weights for all members and $R3$ adjusting the weights only considering type of relationship (family, friends or acquaintances). Results of the Logit model using R, R_2 and R_3 are shown in the next section.

The quality of network appears to be significant with the three types of calculations. Besides, the values of the marginal effects are very similar for both the quality and the size of the network.

3.3 Results

For each individual that participated in the survey, we calculate the quality of his/her network (R) by using equation (2). Figure 4 corresponds to the histograms of the quality for the treatment group (above), the control group and the complete sample. We can observe that the frequencies for the control group are more concentrated below the value 0.5, while the bars above are located mainly to the right of 0.5. Individuals that have continue their studies after secondary school, have in average higher quality of social network.[9]

Once R^q has been calculated for all i in both groups, we proceed to estimate the Logit model (4). Besides the quality of the network, we have to include in the vector X other variables that may influence the decision to continue superior schooling. Recall that at least two of the relevant variables are similar for the two groups, since all individuals belong to the same type of socioeconomically disadvantaged neighborhood and posses around the same level of intellectual ability (high enough to carry on superior studies). We then run a first regression including two variables in the vector X: "size of the network" (m, number of ties) and "number of siblings". Table 3 shows the average quality of the network for both groups, as well as the mean of the independent variables . We observe that individuals in the treatment group have in average a higher number of members in their social network, and 0.5 less siblings than the individuals in the control group.

Results of equation (4) are shown in table 4. Contrary to what we expected,

[9] It is worthy to clarify that the quality of the network is not endogenous i.e. it is not the case that students going to college have a better quality of the network precisely because they go to college. The reason is that we have interviewed students from the first year and asked for their friends and close people, explicitily asking them to exclude those appearing after college.

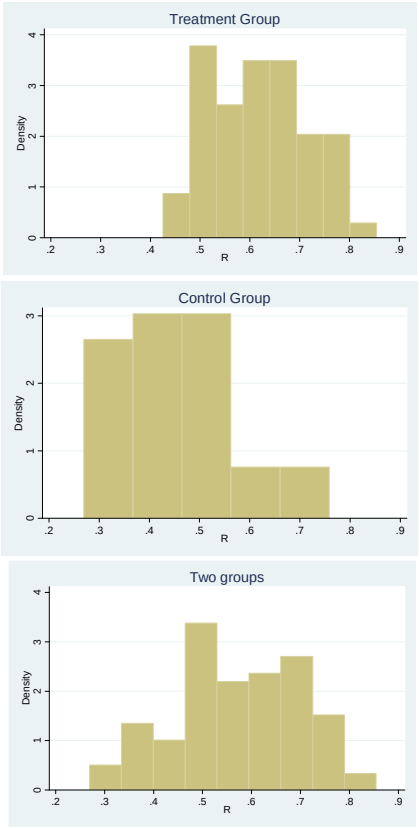

Figure 4: Histogram, quality of the network for the treatment group (above) and the control group.

Table 3: **Mean of the variables**

	R	Size of network	# of siblings
Treatment group	0,6208	24,7	1,7
	(0,9746)	(8,4)	(1,0)
Control group	0,4558	17,1	2,2
	(0,1149)	(6,2)	(1,3)

Source: author calculations based on experimental field work
Standard deviation in parenthesis

the number of siblings is not significant to explain the probability to continue superior studies (columns (2)), which left us with a simpler model whose results appear on the third column of table 4. The coefficient for network quality has the expected sign (positive) and it is significant. The marginal effect indicates that the quality of the network, as expected, does increase the probability of an individual to continue studying after secondary school.

Besides, the size of the network seems to be also relevant to explain schooling decisions. This is expected since a higher number of (relevant) role models allows individuals to compare the situation of their different ties and have more information when forming their expectation on the returns to schooling.

Table 4: **Results of the logit model**

	(1)	(2)		
(R)	14.54732			14.75593
	(3.622724)			(3.610102)
marg eff	1.976019			2.002
($R2$)		13.48816		
		(3.360578)		
marg eff		2.002597		
($R3$)			13.27793	
			(3.323732)	
marg eff			1.950189	
netw size	.1618954	.1616536	.1528159	.161769
	(.055938)	(.0539379)	(.0532306)	(.0553994)
marg eff	.0219909	.0240008	.0224448	.0219567
# siblings*	-.1605066			
	(.3073555)			
marg eff	-.0218022			
_cons	-10.02124	-10.14222	-9.382681	-10.42581
	(2.466409)	(2.303447)	(2.162289)	(2.371418)
LRX_2	50.4	44.8	47.3	50.1
Prob>X_2	0.00	0.00	0.00	0.00
Log lik =	30.15	-32.94	-31.68	-30.29
PseR2	0.455	0.405	0.427	=0.453

* Non significant

Table 4 shows the marginal effects or elasticities of the dependent variable (S) with respect to the quality of the network at different levels. As expected from the Logit specification, marginal effects are higher for intermediate quality levels.

Table 5: **Mean of variables**

Quality	Marginal effect (R)
0.2	0.3031322
0.3	1.156286
0.4	3.043029
0.5	3.391535
Mean (0.5718)	2.002804
0.6	1.480036
0.7	0.4061726

Marginal effects nonsignificant for quality levels higher than 0.7

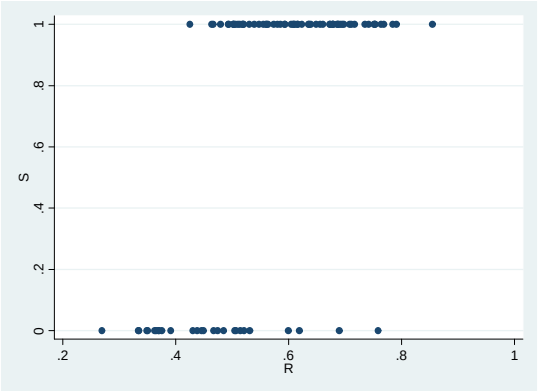

Figure 5: Quality of network and schooling decision (Figure 3 with real data)

3.4 Estimating σ

Figure 5 plots the quality of the network against the schooling status. This is the same figure 3 but with the real data. We are interested to estimate a critical value of R_i^q that represents the network's quality level above which S_i will be most likely equal to 1.

We have claimed that

$$\begin{aligned} S_i &= 1 \text{ if } R_i^q \geq \sigma \text{ or } R_i^q - \sigma \geq 0 \\ &\quad \text{and} \\ S_i &= 0 \text{ if } R_i^q < \sigma \text{ or } R_i^q - \sigma < 0 \end{aligned}$$

Denote $z = R_i^q - \sigma$. Given the symmetry property of the logistic distribution (0.5 is the kink point) we can estimate z by,

$$F(z) = \frac{e^{-z_i}}{1 + e^{-z_i}} = 0.5$$

This is, $z = 0$. Assuming that z is a function of the individual's characteristics included in (4), we may write it as follows

$$z = \hat{\beta}_0 + \hat{\beta}_1 R_i^q + \hat{\beta}_2 \mathbf{X_i'} + \boldsymbol{\varepsilon_i} = 0 \tag{6}$$

Hence, we obtain a value of $\sigma = 0.46$.[10] Individuals belonging to a social network with quality equal or higher than 0.46 are said to belong to a high quality network, and the model would predict that they have continued or will continue their superior studies.

3.5 Key ties

We explore in more detail the information obtained trough the survey, in order to identify 'key ties' such as it is described in definition 3. A key tie is a high quality non family member of the network crucially helping in determining the quality of the social network.

As we have mentioned, it was found that the quality of the network for individuals in the treatment group is in general higher than for individuals in the control group (Figure 4). Likewise, the average quality of the family members is higher. For instance, 21% (27%) of the mothers (fathers) of individuals in the treatment group have university education, while only 8% (13%) of the mothers (fathers) of individuals in the control group got an university degree. This indicates that in several cases, strong ties are playing a decisive role in determining the high quality of the network (parents, siblings, uncles, aunts, cousins).

However, there are also several cases where non relatives play an important role. We found that 20% of individuals have a non-relative key role model (high quality).[11] They have, in general, a family quality below $\sigma = 0.46$. These key ties are, in most of the cases, teachers, friends or family friends. Interaction with people different to their strong ties has apparently widened their life perspectives. A less encapsulated social network allows the potential advantages of positive influence from high quality role models.

4 Conclusions

This paper explores the role of the quality of the network on the schooling decisions of individuals. We want to show that, beyond the income restrictions, individuals with higher quality of network will experience higher probability to continue studying. We compare two groups of people belonging to the same type of poor neighborhood and with similar intellectual ability. The difference between them is that individuals in the first group have continued studying after secondary school, while those in the control group have not continued.

[10] To obtain R we use the mean value of the other dependent variable.

[11] 53% of individuals in the target group have at least one high quality non-relative member of the network with more than one weight adjustment factor applying. However, for only 20% the role of those members was "crucial" as described in definition 3.

We propose a measure for quality network, based on the weighted summatory of the quality of each member of the net. Here we consider that the type of relationship (relative, friends or acquaintances) is not enough criterion to determine the weights of each tie. Other factors such us intimacy, confidence, emotional and economic support are relevant and may convert an initially weak tie into a key tie.

Results of the Logit model corroborates the relevance of the quality of the network in schooling decisions. That is the reason for which we may find individuals belonging to the same community or neighborhood, who share certain attributes like family income and ability, ending up at different equilibrium levels of schooling, expected future income and expected social mobility. If we accept this, a relevant question is, how to strength the quality of networks for individuals in poor neighborhoods? let us mention (not develop) some strategies towards that direction:

- Good students from public schools should be guaranteed a given number of undergraduate scholarships in high quality private universities. This should not reduce the opportunity for them to access to the scarce current scholarship offers. Education in Colombia is polarized, and poor students normally are excluded (not in theory but in practice) from private universities. Students from high quality private universities have much more chances to engage in the labour market, because of good networking and prestigious of the academic programs.

- It is imperative to increase quality in public education at all levels. The current quality gap between public and private education is huge at primary and secondary level-, which implies future disadvantages of students from public institutions to access to superior education, improve the quality of their network, and obtain good positions at the labour market, among others.

- We have seen that, in several cases, the positive influence of a teacher plays an important role in shaping individuals behavior. There should be some mechanisms to strength the relationship teacher-students to exploit this channel of positive influence.

- A more ambitious but highly beneficial proposal is that at least one member of each family should benefit from a scholarship to cover the complete educational cycle. This member will have a positive incidence in the whole family group.

- Currently, there exist the so-called "Godfather" program, through which people in good socioeconomic position helps to support disadvantaged children. Normally it is monetary support, but it would be helpful to extend the support and contact to other levels, so that wider social interaction is promoted.

- Cities must offer more spaces that people from different income groups may indistinctly access. Cities in Colombia are designed to promote polarization. At least from the government initiatives, this practice should be stopped.

- Social housing (viviendas de interés social) should not be constructed in isolated places, as it is usual in Colombia. It is important that they are integrated to the city so that inhabitants may benefit from the infrastructure of the city and may easily socialize.

- If one of the advantages of high quality network is higher information for the labour market, all measures increasing labour information at general level play the same role.

References

[1] Akerlof, G. (1997). Social distance and social decisions. Econometrica Vol. 65 N^o 5. Pages 1005-1027.

[2] Akerlof, G. and Kranton, K. (2002). Identity and schooling: some lessons for the economics of education. Journal of Economic Literature. Vol. XL. Pages 1167-1201.

[3] Anderberg, D. and Anderson, F. (2007). Stratification, Social networks in the labour market, and intergenerational mobility. The Economic Journal 117. Pages 782-812.

[4] Austen-Smith, D. and Fryer, R. (2005). An economic analysis of "acting white". Quarterly Journal of Economics. Vol. 120 N^o 2. Pages 551-583.

[5] Bénabou, R. (1993). Workings of a city: location, education, and production. The Quarterly Journal of Economics Vol. 108 N^o 3. Pages 619-652.

[6] Bénabou, R. (1996). Equity and efficiency in human capital investment: the local connection. Review of Economic Studies 63. Pages 237-264.

[7] Boorman, S. (1975). A combinatorial optimization model for transmission of job information through contact networks. The Bell Journal of Economics Vol. 6 N^o 1. Pages 216-249.

[8] Buhai, S. and Van der Leij, M. (2006). A social network analysis of occupational segregation. Mimeo Tinbergen Institute and Erasmus University.

[9] Calvó-Armengol, A. and Jackson, M. (2004). The effects of social networks on employment and inequality. The American Economic Review Vol. 94 N^o 3. Pages 426-454.

[10] Calvó-Armengol, A. and Jackson, M. (2005). Like father, like son: social networks, human capital investment, and social mobility. Social Science WP 1242. California Institute of Technology.

[11] Castillo, M. and Salazar, B. (2007). Pobreza urbana y exclusión social de los desplazados. Documento de trabajo CIDSE- Universidad del Valle.

[12] Contreras, D. et.al. (2007). The role social networks in the economic opportunities of Bolivian women. Research network working paper N^o R-540. Inter-American Development Bank.

[13] Djebbari, H. (2007). Identification of peer effects through social networks. IZA discussion paper N^o 2652. Université Laval.

[14] Durlauf, S. (1992). A theory of persistent income inequality. Journal of Economic Growth Vol. 1. N^o 1. Pages 75-93.

[15] Goos, M. and Salomons A. (2007). Dangerous liaisons: A social network for the gender wage gap. WP 0722. Kuleuven.

[16] Granovetter, M. (1983). The strength of weak ties: a network theory revisited. Sociological theory Vol. 1. Pages 201-233.

[17] Heckman, J. and Rubinstein, Y. (2001); The importance of non-cognitive skills: lessons from the GED testing program. American Economic Review. Vol. 91 N^o 2. Pages 145-149.

[18] Heifetz, A. and Minelli, E. (2006). Aspiration traps. Discussion Paper 0610. Università degli Studi di Brescia.

[19] Jackson, M. (2003). A survey of models of network formation: stability and efficiency. In: "Group formation in economics: networks, clubs and coalitions". Edited by G. Demange and M. Wooders. Cambridege University Press.

[20] Jackson, M. (2005). The economics of social networks. In: proceedings of the 9th World congress of the econometric society. Cambridge University Press.

[21] Manski, C. (1993). Identification of endogenous social effects: the reflection problem. Review of Economic Studies 60. Pages 531-542.

[22] Marsden, P. and Campbell, K. (1984). Measuring tie strength. Social Forces Vol. 63 N^o 2. Pages 482-502.

[23] Moizeau, F. et. al. (2004). Neighborhood informational effects and the organization of the city. University Paris I.

[24] Oreopoulos, P. (2003). The long-run consequences of living in a poor neighborhood. The Quarterly Journal of Economics. Vol. 118 N^o 4. Pages 1533 – 1575.

[25] Roemer (1993).

[26] Ramírez, M. (2007). Seguimiento de la educación en la Republica de Colombia ¿en que fallamos? Borradores de Economía – Banco de la República.

[27] Requena, F. (2003). Análisis de redes sociales. Orígenes, teorías y aplicaciones. Centro de Investigaciones Sociológicas. Siglo XXI de España editores S.A.

[28] Streufert, P. (2000). The effect of underclass social isolation on schooling choice. Journal of Public Economic Theory Vol. 2 N^o 4. Pages 461-482.

[29] Vivas, H. (2007). Educación, background familiar y calidad de los entornos locales en Colombia. Doctoral Thesis, Department of Applied Economics. University of Barcelona.

[30] Yamauchi, F. (2005). Social Learning, neighborhood effects, and investment in human capital: evidence from Green-revolution India. International Food Policy Research. WP 190.

A The survey

Table 6: **Survey**

	1	2	3	4	5	6	7	8	9	10	11
Father											
Mother											
Sibiling 1											
...											
Cousin 1											
...											
Uncle/aunt 1											
...											
relative in law 1											
...											
Godmother											
Godfather											
Friend 1											
....											
Family friend 1											
...											
boyfriend/girlfriend 1											
...											
School Teacher 1											
...											
Neighbor 1											
...											
Group or asoc. 1											
...											
Other 1											
...											

1: Gender
2: Age
3: Level and year of education
4: Occupation
5: Frequency of Contact
6: Type of contact
7: Economic support
8: You ask for advice
9: They give you advice
10: Emotional support
11: Admiration