# The impact of social capital on children educational outcomes: the case of Tanzania

Y. Baende Bofota

Discussion Paper 2013-3

Institut de Recherches Économiques et Sociales
de l'Université catholique de Louvain


UCL

# The impact of social capital on children educational outcomes: The case of Tanzania

Youyou Baende Bofota*

**Abstract**

This paper presents an empirical analysis of the relationship between social capital and children' educational outcomes in Tanzania, using panel data from the Kagera Health and Development Survey (KHDS). By exploiting the panel structure of the data, we use several econometric techniques - fixed effect, first difference and 2SLS - to address social capital endogeneity issue and omitted variable bias. We find evidence that social capital available in the family affects significantly student attainment and that the magnitudes are large enough to explain a substantial proportion of variation in children schooling in Tanzania in the short term. More importantly, this positive impact lasts over the long term.

**Keywords:** Social capital, Education, Developing countries, Tanzania.

**JEL Classification:** 010, 012, 125

§ I am especially grateful to William Parienté for very helpful comments and guidance.

*PHD. IRES, Université Catholique de Louvain, Belgium. E-mail address: youyou_baende@yahoo.com

# 1 Introduction

Since the late 1990s a considerable amount of research has been devoted to the connection between social capital and education.

According to the economic literature, three broad characteristics emerged as the main determinant of educational outcomes. The characteristics identified were: (1) child characteristics such as sex, age, biological child of the head of the household, parent schooling (Card, 1999; Beegle and Burke, 2004). (2) Also, features of the community including the availability of schools in the community, distance to school and elements of school quality like student to teacher ratio, number of teachers in school, number of blackboards and availability of textbooks are among those determinants (Glewwe and Jacoby, 1994; Bommier and Lambert, 2000; Ainswoth *and al.*, 2003; Beegle and Burke, 2004). (3) Household characteristics such as years of schooling of the head of the household, size of the household, the amount of family income or household resources were also included (Haveman and Wolfe, 1995; Lloyd and Blanc, 1996; Thomas *et al.*, 2003; Beegle and Burke, 2004).

Focusing particularly on household determinants of children education, a large body of empirical researches persuasively show that households characteristics such as parental financial and human capital are important determinants of schooling. Notwithstanding these empirical evidences, there still remains a substantial proportion of variation in children schooling that remains unexplained. Researchers estimated to less than one-half of the variance in years of schooling obtained (a measure of educational attainment) that can be accounted for by parental background (Tatchman *et al.* 1997, Duru-Bellat, 2004).

Coleman (1988) important contribution on the impact of social capital on education, demonstrates that parent human and financial capital are made available to children through their family social capital. Since Coleman's seminal contribution, social capital has been accepted by numerous authors as a key factor that impact positively children educational outcomes (see, for example, Carbonaro, 1998; Coleman, 1988; Furstenberg and Hughes, 1995; Teachman *et al.*, 1996, 1997; Hofferth *et al.*, 1999; Morgan and Sorensen, 1999; Bryk and Schneider, 2002; Bortnick, 2004; Sandefur *et al.*, 2006; Bassani, 2007a).

For instance, Tatchman *et al.* (1997) argued that, Coleman does not envision social capital as simply another variable affecting children schooling. They underline that while a marginal effect may be anticipated, social capital sets the context within which the financial and human resources possessed by parents may impact schooling. Likewise, findings across studies highlight the effects of social capital on various educational outcomes, especially educational achievement; e.g., grades, grade point average and test scores (Carbonaro, 1998; Hao and Bonstead-Bruns, 1998;

Lopez, 1996; Pong, 1998; Sun, 1998; Israel *et al.* 2001; Morgan and Sorensen 1999; Muller 2001; Roscigno and Ainsworth-Darnell 1999; McNeal, 1999; Muller, 2001); and educational attainment; e.g., high school completion, high school graduation, college enrolment, years of schooling and dropping out (Carbonaro, 1998; Teachman *et al.*, 1996, Hofferth, Boisjoly, and Duncan 1998; Wilson, 1999; Muller and Ellison, 2001; Kim and Schneider, 2007). Still others analyze the effect of social capital in educational attendance (Grootaert, 1999).

Thus, any discussion around the determinants of children educational outcomes, must consider social capital within family as an important factor. In other words, improving educational outcomes requires, among other things, paying attention to the available social capital in the family.

Social capital has been defined in a variety of ways. Social capital concerns the quantity and the quality of relationships and the value of the resources available through these relationships (Putnam, 2000). Put differently, it consists of resources embedded in social relations which can be mobilized to increase the likelihood of success in a purposive action (Bourdieu, 1983; Coleman, 1994; Lin, 2001b). The key insight is that albeit beneficial for the individuals, social capital is not an individual characteristic or a personality trait but a resource that resides in the networks and groups to which people belong (Mouw, 2006).

Coleman used indicators such as family structure, parent-child discussions, and intergenerational closure to measure social capital that affects education. He particularly emphasizes the importance of education oriented relationships between the parents and the child in improving children educational outcomes. Following this an extensive empirical literature offers evidences that family structure may create better outcomes for children. This empirical evidence particularly shows that the presence of two parents at home, a lower number of siblings, higher parental educational expectations, parental monitoring, parent-children connections, and parents' communication with their children's school teachers and friends impact positively educational outcomes.

However, recent empirical researches have refined Coleman measures and explore new routes about the influence of social capital on students' educational performance beyond Coleman measures. These studies, while using the Coleman's social capital framework, departed from Coleman's operationalization and consider the effects of "non educational oriented relationships" of parents on children education. Dika and Singh (2002) excellent review of the empirical literature on social capital and education in the period 1990 to 2001 argues that several authors used very different indicators to measure social capital.

For instance, Hofferth *et al.* (1995) measure social capital by the access to help from friends arguing that it promotes college attendance. Bankston and Zhou (1995) looked at literacy skills and cultural identity as sources of community social capital for students, while Stanton-Salazar

and Dornbusch (1995) consider language proficiency as a source of cultural and social capital. Using Lin (1990)'s social resource theory, Stanton-Salazar and Dornbusch (1995) measured social capital with social network indicators such as status of network member and number of non-kin. All these measures have been proved to affect positively various children educational outcomes.

Recently, Buchel and Duncan (1998) greatly expand Coleman's measures of social capital in a study of German households. This study was conducted to investigate whether parental social activities, such as attending cultural events, doing volunteer work, participating in active sports, socializing with friends relatives, or neighbors and helping friends, relatives and neighbors may affect educational attainment of children (Buchel and Duncan, 1998). This research concluded that these parental activities constitute a positive form of social capital for the children and reported significant linkages between some of the fathers' activities, such as sports involvement and volunteer work, and children' educational attainment (Buchel and Duncan, 1998).

Most of the studies cited above employ extensively OLS estimation which may be problematic in the presence of omitted variables and endogeneity. Indeed, tests for the existence of social capital effects typically examine the sign and significance of the coefficient of social capital in the regression analysis. So far, researchers that have used simple regression estimation end up with very cautious statement about the presence of social capital effects, avoiding a definitive conclusion about the causal effects of social capital on educational outcomes. Many others researchers have used multilevel models to analyze social capital effects on children educational outcomes (Israel *et al.* 2001; Horvat *et al.* 2003; Goddard, 2003). Yet, this technique does not address biases from omitted variables or endogeneity of social capital.

Thus, a number of papers questioned whether it is possible to rule out the fact that the results for the social capital effects on schooling were not driven by a selection bias, reverse causality or by some unobservable factors not identified by the researchers (Lee and Bryk, 1988).

Moffitt (2001), Manski (1993, 2000), Durlauf (2002), Durlauf and Fafchamps (2004), and Mouw (2006) all depict the difficulties of estimating social capital in the presence of unobservable characteristics and endogeneity. Some important lessons can be drawn from their analyzes. Firstly, in the presence of a bias in the social capital model, caused by fixed, unobserved characteristics, then a first difference or fixed effects model using longitudinal data can provide an unbiased estimate of social capital (Mouw, 2006). Example of these studies are among others, Morgan and Sorensen (1999); Angrist and Lang (2002); Hanushek *et al.* (2003) and Arcidiacono and Nicholson (2005). They all used fixed effects models to control unobserved factors at the individual or school level. These researches provide only mixed results for social capital effects on education. Secondly, in the presence of endogeneity, authors such as Evans *et al.* (1992); Neal (1997): Hao and Bonstead-Bruns (1998); Cipollone and Rosolia (2007) used instrumental

variables regressions to correct for endogeneity of social capital.

In general, recent theoretical developments on this topic use instrumental variables to account for the endogeneity of social capital. However, in most of the cases, there is a lack of the necessary data to make clean tests and to reliably account for potential omitted variables bias and endogeneity. Thus, there is a need of empirical relevance of social capital in educational outcomes that may ground effective policies.

In this paper, we conceptualize social capital at the family level and, using panel data, we analyze the short and long term impacts of family social capital on educational attainment in Tanzania. We investigate these effects by controlling for heterogeneity across gender, child position in the household (biological vs foster children) and age groups. Family's social capital may impact educational outcomes through several mechanisms. One of the most important mechanisms is that parents' connections to resourceful persons can provide access to important resources such as information, money, control, and potentially helpful contacts for children.

Our research may yield an interesting contribution to the empirical literature on social capital and educational outcomes which has paid insufficient attention on developing countries. This absence of researches is mainly due to the lack of data on social capital covering this region. Since social capital is an important factor in education performance, its shortage may be detrimental to disadvantaged students, especially in developing countries.

This study considers and empirically determines that household-level social capital is an important factor that impacts children educational outcomes. Our empirical analysis examines the years of schooling as the schooling outcome. Our results suggest that children living in households with a higher level of social capital get more education in the short term as well in the long term.

The remainder of this paper is organized as follows. Section 2 presents the data and variables used in the analysis, and discusses the results obtained under the various specifications. Section 3 features heterogeneity tests. Section 4 concludes.

# 2 Data description

## 2.1 Setting

Formal education in Tanzania was initially modeled on the British system. Formal schools were associated with socioeconomic status as graduates were able to gain higher waged employment in government service (Jung and Thorbecke, 2003).

In Tanzania, prior to independence in 1961, access to basic education was limited and un-

equally distributed by region and gender (Al-samarrai and Pessgood, 1998). For instance, rural Tanzanians were left at a severe disadvantage for educational opportunities during the colonial period. At the end of this period, nearly 80 percents of adults in Tanzania were illiterate (Cameron, 1980). When Tanzania gained its independence in 1961, the country initiated major education reforms whose objective was to deliver an education that was both relevant to individual communities and to foster a sense of national identity (Mushi, 2009). Despite these reforms, the education system was hit by several problems. Schools were characterized by severe shortages of teachers, approximately 10 teachers were staffed in a typical primary school in Kagera during the survey periods. Educational tools such as text books, and infrastructure such as classrooms or sanitation facilities were not always available. The rural schools faced particularly many difficulties because of high student-to-classroom ratios at 180 to 1 and teacher-to-student ratios at 72 to 1 from 1974 to 1991 (Sifuna, 2007).

In the beginning of the last decade, Tanzania had a primary school attendance rate that was much lower than the average of the sub-Saharan Africa (Bommier and Lambert, 2000). Also, enrollment rate for secondary schools in Tanzania is among the lowest in the world (Ainsworth *et al.*, 2002). By the end of the twentieth century the level of quality of the education system had reached crisis point (Kuleana, 1999; Galabawa *et al.*, 2000; Lwaitama *et al.*, 2001).

## 2.2 Data

Our empirical work consists of relating children educational outcomes to measures of household social capital and more traditional covariates at the child, family and community level that may affect children education. For that purpose, we exploit data from the Kagera region in Tanzania. The Kagera Health and Development Survey (KHDS) was part of a research project conducted by the World Bank and the University of Dar es Salaam (Burke and Beegle, 2004; Ainsworth *et al.*, 2002; Beegle, 2005). The KHDS contains 5 waves as it surveyed individuals during 1991-1994 (wave 1 to wave 4) and follows up in 2004 (wave 5). Wave 1 and wave 5 are annual surveys and asked households about the past 12 months. On the other hand, wave 2, 3, and 4 are half-year surveys and asked households about the past 6 months. The survey was originally designed to measure the impact of prime-age adult deaths and illness due to HIV/AIDS on the welfare outcomes, including child schooling, nutrition, adult labor force participation and poverty status. The data cover also detailed information on consumption, expenditure, business activities, education and health status, as well as community-level information on health, price and education systems.

Over 816 households were surveyed in the first wave. After each round, some households

dropped out and were replaced, so that the total number of households was 915 (totalling to 6,204 individuals) (Ainsworth *et al.*, 2004 and Beegle *et al.* 2006b). The 2004 survey attempted to resurvey all individuals who had been interviewed during any of the four rounds during 1991-1994 (KHDS-1) survey. The new households formed from these earlier respondents were interviewed so that, starting from the 915 original households, the 2004 (KHDS-2) sample consisted of 2700 households (See Beegle *et al.* (2006) for the follow-up survey). The Tanzania-Kagera survey has a low attrition rate of around 10 percent per wave. The main reason for attrition is death in the household and migration, which led to the relocation of the household.

### 2.2.1 Sample design

Our analyze proceeds in two steps using two different panels covering two different time periods. Firstly, using two rounds of KHDS-1 survey, i.e. wave 2 in 1992 and wave 4 in 1994, we construct a child-level two waves panel data, in what we label panel 1. We use the 1992 and 1994 waves because, as discussed above, the KHDS survey has an average interval between waves of six to seven months and most of our key variables (eg. age of child, grade...) likely do not change much in 6 months. Thus, we select waves, with at least 1 year gap to allow some variability in data, necessary to perform a fixed effect analysis[1].
Social capital lagged one period, can be used to establish the causal relationship[2] in our instrumental variable regression.

Panel 1 follows children aged 7 to 16 years old and allows to testing the effects of social capital in the short run. In Tanzania, school begins at seven and primary school is compulsory. The Tanzanian education system consists of a seven year primary school cycle, followed by four years of lower secondary and two years of upper secondary school[3]. The choice of the upper age bound (16 years old) is justified by the importance of enrollment delay, grade repetitions and the high prevalence of child labor in Kagera. So, we do not expect a child to start school after 16 years old. We exclude any child in this age range who is identified as non-relative of the household head, such as boarders, servants and tenants. These restrictions result in a sample size of 1,928

[1]The only choice we had in the short term analysis was the combination of wave 2 and 4 (labeled panel 1) , wave 2 to 3 having only 6 months or interval, as wave 3 to wave 4. Also, the rate of attrition between wave one and wave four is more than 30 percent, which not only reduces the sample size, but also raises the issue of sample selection

[2]As we suspect that our social capital variables may be endogenous and that finding external instrument for social capital is difficult, we use the realization of social capital in waves one and three as instrument for the 2SLS regression.

[3]Primary schooling in Tanzania is completed after grade seven, lower secondary schooling after grade 11 and upper secondary schooling after grade 13 (Ainsworth *et al.*, 2005).

children in 522 households across 50 villages.

Secondly, the analysis is moved forward, over a 10-year horizon, combining two-waves panel dataset (labelled panel 2) consisting of the last wave of KHDS-1, completed in 1994 and KHDS-2 in 2004. There is no other survey between 1994 and 2004. Setting up the dataset in this way allows for testing the effects of social capital in the the long run, since we are following the same individuals ten years after.

To study the long-term effect of social capital on children's education, we also include children aged between 0 to 6 years old in 1994 in our panel 2 sample. Importantly, the follow up survey in 2004 provides information on these children educational outcomes in 2004 when they are aged between 10 and 19 years old. Thus, Panel 2 follows children aged 0-16 years at wave 4 - reinterviewed in wave 5 - rather than the 7-16 year target group of panel 1.

The panel 2 sample is therefore restricted to those individuals interviewed in wave 1994 and in wave 2004. This panel contains 3,418 individuals in 626 households across 51 villages. Table 3 and 4 provide the descriptive statistics of the variables used in the study. The dependent variable is years of schooling, which ranged from approximately 1 to 11 years in panel 1, and 1 to 18 in panel 2.

## 2.3 Variables

We discuss the most salient details of the most important variables below, and other variables will be listed in Appendix A.

### 2.3.1 The dependent variable

Our dependent variable is school attainment. School attainment was measured by years of schooling attained as we wish to capture the quantity of human capital accumulated. Years of schooling[4] in this study do not necessarily indicate the total years of education achieved by the children, especially if they are still going to school. In our two panels, most of children are still enrolled in school and may complete additional years of education.

Three types of schools are identified in the sample: Public, Secular and Religious schools. Attendance levels vary considerably among schools, and are often high in public school than in other types of school.

[4]Throughout the rest of this paper, we use the terms years of schooling, education and educational outcomes interchangeably

## 2.4 Independent variables

### 2.4.1 Social capital

Family social capital is the key independent variable in our analysis. We use four indicators to proxy social capital in the KHDS data: household relative's network, household friends and neighbors network, the amount of remittances received and whether both parents are present at home. We consider, following the literature on social capital that these variables capture social capital and even though they may not be very good proxies for social capital given the data at hand, a common denominator of them may stand as an indicator of social capital.

**(a) Household's relative network.**

Social capital is approximated firstly by the size of the family's relatives network. The household questionnaire in the KHDS data provides information on relatives to which household members are linked, including non-resident parents, family's children living elsewhere and close relatives. This gives the possibility to infer family network variable, thus to measure what we label relative network. Using the same data, De Weerdt (2009) measures the network dimension of social capital in Tanzania by the number of family's children living outside of the household. Indeed, he argued that contacts with children living elsewhere build up family networks (De Weerdt, 2009).

**(b) Households friends and neighbor's network.**

We combined every occurrence of friends and neighbors in relationship with a given household in one category labeled friend network, which constitutes a dimension of social capital in its own. The social capital literature used extensively such proxy to measure social capital. Indeed, Putnam (2000) defines social capital as "connections among individuals-social networks and the norms of reciprocity and trustworthiness that arise from them". This includes also interactions with friends and neighbors. Li *et al.* (2005), using data from three waves of the British Household Panel Survey, developed measures of neighborhood attachment (interactions with people in neighborhood) and social network (presence of social support networks in the form of intimate relations) as indicators of social capital.

On average, a household in our sample interact with 3.5 persons: 2.1 of them are identified as relatives and 1.4 of them as friends or neighbors. Notice that, Tables 4.1 and 4.2 shows that these networks measures have important maximum values. For instance, the maximum number

of relative is 18 and that of friends is 12 in panel 1 data.

### (c) Remittances received

Receiving remittances may constitute a beneficial form of social capital and is not only a mean of accessing resources. It implies that the household maintains relationships with relatives, friends and neighbors, and potentially has large migration network (Becker, 1988; Gubert, 2002). Previous studies in Kagera conceptualize remittances as a form of social capital. Using the same data, Litchfield and McGregor, (2008) measure social capital in Tanzania by the value of remittances received to infer on the ability of households to call on external resources via networks. Goldin and Katz (1999) argue that the expansion of higher secondary education in the United States before World War II was influenced to a considerable extent by social capital. They measure social capital by the resources allocated by local communities to primary schooling. The literature considers that the potential access to gifts and loans of money or time assistance through one's contacts and relationships represents the stock of social capital (Hofferth *et al.*, 1998). This observation is reminiscent of the contention about the strength of relationships in Kagera. We then use as a measure of the strength of relationships the log of the remittances received from friend and relatives.

### (d) Two Parents family

Coleman's (1988) definition of social capital includes relationships within the family and outside of it. Coleman argues that family structure is an important dimension of social capital. It indicates a structural deficiency in the creation and maintenance of social capital. Specifically, one-parent families have less time to devote to parent-child interactions than two-parent families. Accordingly, children belonging to single-parent and other types of family structure are associated with lower levels of social capital than children in two parent families (Coleman, 1988). Family structure in our sample was measured with three dummy variables: single-parent family, two parent's family, and other family structure. Households where two parents were present were categorized as a two parent family which was consider as an indicator of social capital[5]. The panel 1 sample shows that the largest percent (41.6%) of children reside in a two parent family while 32.5% were living in a household with a single-parent family. In panel 2, still 23.96% were living in a two-parent family. Panel 2 has a much lower percentage of two- parents family

[5]Israel, Beaulieu and Hartless, (2001) use among other measures of social capital the number of parents present in the household

because during the decade, many of the children interviewed in 1994 formed their new, separate households.

### (e) Social capital index

The first three proxies of social capital were subjected to Principal Components Analysis (PCA)[67]. PCA is a multivariate statistical technique used to reduce the number of variables in a data set to a lower dimension to reveal the simplified structures that underlie it. Precisely, from an initial set of $n$ correlated variables, PCA creates uncorrelated indices or components, where each component is a linear weighted combination of the initial variables. For example, from a set of variables $X1$ through $Xn$,

$$\begin{aligned} PC_1 &= b_{11} * X_1 + b_{12} * X_2 + \cdots + b_{1n} * X_n \\ \vdots \quad & \quad \vdots \quad \vdots \quad \cdots \quad \vdots \\ PC_m &= b_{m1} * X_1 + b_{m2} * X_2 + \cdots + b_{mn} * X_n \end{aligned}$$

where $b_{mn}$ represents the weight for the mth principal component and the $nth$ variable. This index was calculated to measure social capital at household level. We save the first principal component as $sk1$, which explains about 44 percent of the total variation.

Figure 4.1 shows the distribution of the social capital index to get a clear picture.

[6]For the sake of tractability, the variable capturing the presence of two parents in the household is used in a separate regression.

[7]The index contains positive and negative values indicating the negative or positive correlation of a variable with the component.

Figure 1: Distribution of the social capital index

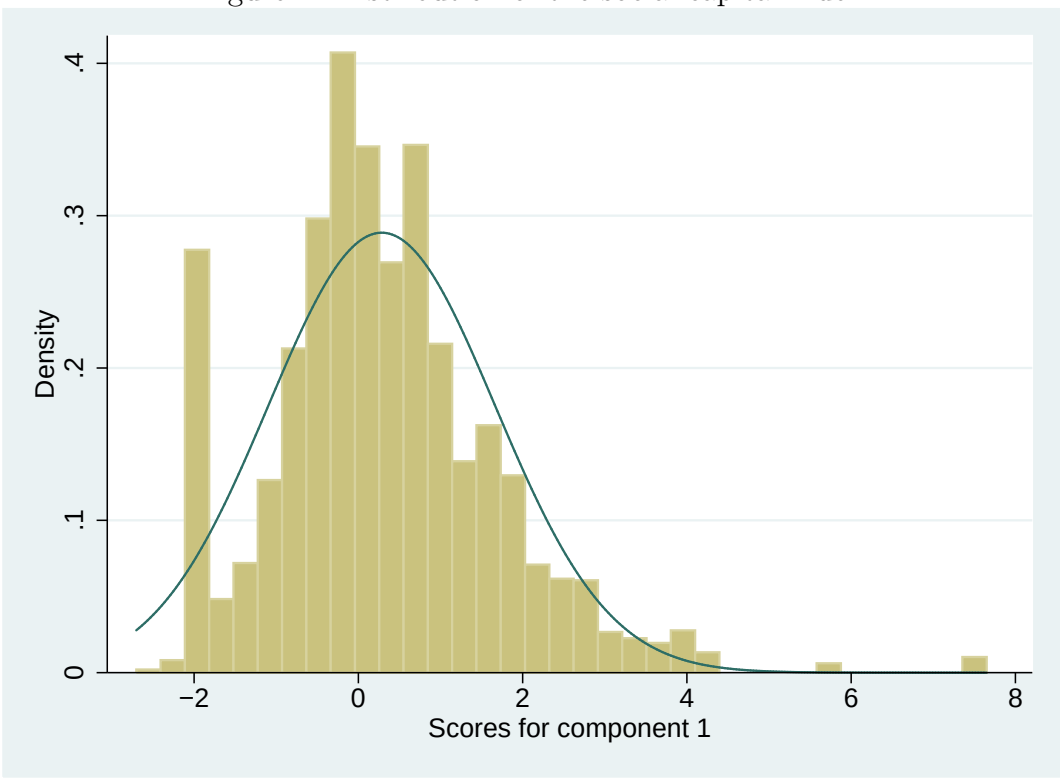


Table 1: Summary statistics Social capital 1992-1994

| Variable | Mean | Std. Dev. | Min. | Max. | N |
|---|---|---|---|---|---|
| Relatives | 2.114 | 2.687 | 0 | 18 | 1928 |
| Friends and Neigh | 1.39 | 1.841 | 0 | 12 | 1928 |
| Remittances | 5.366 | 2.868 | 0 | 13.642 | 1926 |
| Two Parents | 1.156 | 0.807 | 0 | 2 | 1928 |

Table 2: Summary statistics Social capital 1994-2004

| Variable | Mean | Std. Dev. | Min. | Max. | N |
|---|---|---|---|---|---|
| Relatives | 5.58 | 5.245 | 0 | 32 | 3262 |
| Friends and Neigh | 2.545 | 2.380 | 0 | 12 | 3210 |
| Remittances | 7.348 | 3.667 | 0 | 27.408 | 3303 |
| Two Parents | 0.241 | 0.427 | 0 | 2 | 3303 |

Both waves contain detailed information on the elements of social capital index so that we may compare the evolution of social capital between the two panels in table 4.1 and 4.2. Paired t-test reveals significant differences in social capital indicators between the two waves. The number of close relatives and the amount of remittances received has almost doubled between the two waves. In 1994, an household has on average 3 relatives relationships, while the average number of relatives is almost 7 in 2004. The maximum number of friends and neighbors reported is 12 in both waves. Turning to the remittances received, they increased by 14% from 1994 to 2004.

## 2.5 Control variables.

In addition to social capital variables, we include a variety of additional explanatory variables to explain children's educational attainment. We control for child and household characteristics that may affect children educational outcomes. Following the literature on educational achievement and attainment, we include as child characteristics, the age of children and whether the child was at home while attending school.

The KHDS survey contains information on parents and household characteristics. Parental characteristics might affect children's educational outcomes through various channels. First, the financial resources within the family may affect the investment in children education, particularity in the presence of imperfect capital markets (Rosenzweig and Wolpin, 1993). We measure financial resources of the family by the log of per capita expenditures. Expenditures can serve

as a proxy for the financial situation of the household the individual grew up in.

Beyond these financial aspects, The KHDS data gather information on parent and head education. The majority of parents in our sample has no formal schooling and we measure father and mother education using dummy variables, where the dummy variables take a value one if the mother (father) has at least one grade of formal schooling, zero otherwise. The model allows school factors that are indicative of the quality of school such as the number of teachers to influence educational attainment of the child. We also include variables measuring the distance to school because Bommier and Lambert, (2000) found that distance to school in Tanzania had a significantly negative effect on years of schooling. However, this variable was dropped as it does not enter significantly in any of our estimations.

## 2.6 Descriptive statistics

As said before, enrollment is compulsory in Tanzania, so that the problem to deal with when examining schooling patterns is delayed enrollment. Examining the two panels, it appears that 94.24% of the children interviewed in the panel 1 (wave 1992-1994), reported to be enrolled in school but only 55% of them were attending school in the due time. We have a better pattern for panel 2 (wave 1994-2004) where roughly 70.4% were attending school in the due time. This reveals the often cited difference between "enrolment" and "effective school participation".
If being enrolled is the condition to attend school, the above figures show clearly that in our sample, roughly 45% and 30% respectively for the panel 1992-1994 and 1994-2004 of children aged 7-16 enrolled at the beginning of the year, actually, do not attended school. Further, almost 27.42% of the children in the sample, although enrolled to school have never attended school by the age of 9 and 20% was in their first grade of primary school by the age of 10. Still 6% of children aged 12 reporting not having completed at least the first school grade. Figure 4.2 in appendix A that plots years of education against different age groups confirms these trend.

The higher percentage of delayed enrollment suggests that this seems to be an important phenomenon in Tanzania. There are various reasons causing delayed enrollment such as child labor, poor health, inter-term vacations, parental attitude of whether their child is ready to go to school, a perceived low return to school (...). This phenomenon has been also noted in many others low income countries (see Glewwe and Jacoby, 1994).

We did not observe any differences in enrollment or delaying patterns between girls and boys in the two panels. This latter finding is consistent with that of Skipper and Simonsen (2007).

Turning in the characteristics of the household in panel 1, table 4.3 reports that 27% of the household heads are female; their average age and years of schooling are respectively 51.2 and 5.1. The average household size is 7.6 individuals. Most of the fathers (68.86%) and mothers (62.74%) have not completed primary school.

Table 3: Summary statistics 1992-1994

| **Variable** | **Mean** | **Std. Dev.** | **Min.** | **Max.** | **N** |
|---|---|---|---|---|---|
| Years of education | 3.602 | 2.023 | 1 | 11 | 1538 |
| Grade | 2.901 | 2.309 | 0 | 11 | 1910 |
| Age | 12.898 | 2.722 | 6 | 23 | 1928 |
| Social capital index | 9.410 | 1.241 | -1.879 | 5.693 | 1926 |
| Household size | 7.686 | 3.478 | 1 | 26 | 1928 |
| Mother education | 4.182 | 2.64 | 0 | 11 | 944 |
| Father education | 5.233 | 2.557 | 0 | 19 | 1001 |
| Grade of the head | 5.052 | 2.762 | 0 | 19 | 1703 |
| Distance to school | 17.714 | 76.143 | 0.1 | 800 | 1832 |
| log Expendt | 10.182 | 0.703 | 7.848 | 13.893 | 1928 |
| Number of teachers | 10.485 | 3.699 | 4 | 23 | 1896 |

Table 4.4 shows selected means for our 1994-2004 sample, where children are aged between 10 and 30 in 2004[8].Individuals included in this sample have an average of 5.7 years of education, 42.9% of them have completed or are attending primary school in wave 5 (2004).

It is noteworthy that, over the decade some variables exhibit negligible change. Inspection of table 4.4 shows negligible change in the parent level of education level between the two waves.

[8]the upper bound (30 years old) may be due on misreporting of age in 1994 were we selected 0-16 years.

Table 4: Summary statistics 1994-2004

| Variable | Mean | Std. Dev. | Min. | Max. | N |
|---|---|---|---|---|---|
| Grade | 3.092 | 3.185 | 0 | 18 | 3404 |
| Age | 12.463 | 7.217 | 0 | 30 | 3418 |
| Social capital index | 0.284 | 1.381 | -2.704 | 7.650 | 3288 |
| Household size | 7.917 | 3.552 | 1 | 26 | 3418 |
| Mother education | 4.782 | 2.765 | 0 | 11 | 872 |
| Father education | 5.354 | 2.495 | 0 | 18 | 904 |
| Grade of the head | 5.121 | 2.698 | 0 | 17 | 1708 |
| Distance to school | 30.409 | 145.881 | 0 | 2500 | 2610 |
| log Expendt | 10.631 | 1.066 | 0 | 14.282 | 3418 |
| Number of teachers | 10.975 | 4.307 | 4 | 29 | 3405 |

## 2.7 Attrition 1994-2004

Out of the 1,984 children from the initial 1994 sample, a total of 1,319 children, i.e. 66 % of them were found in the original villages and reinterviewed. 159 children (8.0%) were traced wherever they were living and reinterviewed. Information was lost only for 506 (25.5%) of them, either because they were tracked but not found (233 children, 11.7% of the sample), or they were missing (179 individuals, 9% of the sample), or because they died (a total of 94 children, 4.7% of them died).

We conduct t-tests to compare the attriters and non attriters along observables dimensions (see tables 4.16, 4.17 and 4.18 in appendix D). We do this first by comparing the means of selected child as well family characteristics in both samples to see if they differ. Secondly, we test if the level of social capital between the two groups differs. These tests indicate that attrition is mostly age and grade related. Hence, non attriters are different from attriters when considering children background such as age and the highest grade attained. There are apparently no differences when we consider family backgrounds in terms of expenditure and parental education between the two groups.

As much of the attrition in our sample is associated with observable demographic changes, we infer that it could be that the presence of time-invariant unobservables may likely explain this attrition. The fixed effect estimation strategy used here removes all sources of time-invariant unobservables while addressing this potential source of attrition bias[9].

[9]In appendix D, attrition is regressed upon child and household characteristics which include among others, measure of age of the child, mother's schooling, and social capital variable. The regression results for 1994-2004

# 3 Estimation strategy

We follow two separate but interlocking strategies to analyze the impact of social capital on education. We analyze both the short-term and the long-term impacts of social capital on educational attainment controlling for child's, family and community characteristics. Our fundamental idea is that by controlling for those characteristics, we should be able to determine the extent to which educational successes are attributable to social capital available in the family.

## 3.1 Short run regression

The model estimated in the short term assumes that social capital affects contemporaneously educational attainment and considers years of schooling as the educational outcome.

For child $i$ in household $j$ in village $k$ at year $t$, we may estimate the following econometric model:

$$E_{kjit} = \alpha_0 + X_{kjit}'\beta + \gamma S_{kjt} + W_{kjt}'\rho + Z_{kt}'\lambda + \epsilon_{kjit} \quad (1)$$

Where $i$, $j$, $k$ index individual level, household level and community level variables respectively. We have two periods : $t = 1992$ and $t = 1994$, $E_{kjit}$ is the outcome and $S_{kjt}$ the social capital index. $X_{kjit}$ is a vector of child characteristics, $W_{kjt}$ a vector of other household characteristics, $Z_{kt}$, a vector of community characteristics and $\epsilon_{kjit}$ is the error term.

The parameter $\gamma$ measures the impact of social capital on children education. If social capital matters, then $\gamma$ should be positive and significant. But, specification (1) may entail many problems that may result in biased parameters.

The first problem with this econometric specification in (1) is the potential endogeneity of the current period social capital index. Specifically, some of the components of social capital are potentially endogenous. For example, it is well known in the social capital literature that building networks can be an option under the control of individuals.

Measurement errors are also common in most studies of social capital. Not only good proxies to measure social capital were not available in our data, but it is also difficult to correctly count a family number of friends and relatives. We used in this study self-reported information given by the respondents which may be affected by their imperfect recall and tendency to report rounded numbers. Thus, we suspect that there are some measurement errors in our networks variables.

is reported in table 4.18. They indicate that attrition is negatively associated with age but unrelated to social capital

The second problem is related to child specific time-invariant unobservables such as child's innate ability to perform well in school. Given these unobservables, our estimation strategy must be sensitive to violations of the assumption of zero correlation between social capital and the error term.

OLS may not adequately address the important problems just mentioned. If endogeneity and heterogeneity are presents, then the social capital parameter estimated and its standard error estimator are biased. Unless we find enough information on a child's background or suitable instruments for predicting social capital, then the interpretation on this coefficient should be considered carefully, since it will be biased. We then focus on investigating the relationship between social capital and education in a more causal framework, estimating a child fixed effects model, which controls for unobservable heterogeneity that may bias the OLS estimate. In addition, instrumental variable (2SLS) approach is employed to control for potential endogeneity of social capital. Technically, the 2SLS estimator is a very appropriate method and more efficient than OLS to deal with the problem of endogeneity.

Let us suppose that the child innate ability to perform well in school $c_{kji}$, may affect its schooling outcomes, then we can include $c_{kji}$ into the model (1) in order to account for this unobserved factor. We then expand equation (1) into:

$$E_{kjit} = \alpha_0 + X_{kjit}{}'\beta + \gamma S_{kjt} + W_{kjt}{}'\rho + Z_{kt}{}'\lambda + c_{kji} + u_{kjit} \quad (2)$$

In this model, we denote the time-invariant, unobserved child characteristics $c_{kji}$ and separate it from the original error term, $\epsilon_{kjit}$, with $\epsilon_{kjit} = c_{kji} + u_{kjit}$. Hence, $c_{kji}$ represents the child fixed effect and model (2) can be estimated using fixed effects method. This method controls for all time invariant factors that affect a child's schooling. Yet, the fixed effect estimator imposes a strict exogeneity assumption on the regressors within the model.

Strict exogeneity requires that the error term in equation (2) be unrelated to past, present, and future values of social capital. Indeed, social capital is strictly exogenous if $E[S_{kjt}u_{kjis}] = 0$ for all $t$ and $s$.

We handle carefully this issue and assume that the assumption of strict exogeneity of social capital could be challenged in our analysis. We cannot consider a variable to be both endogenous and exogenous in the short term. In general in the short term, current period $S_{kjt}$ is correlated with $u_{kjit}$. To deal with this endogeneity, we calculate first differences on both sides of equation (2) which can be written as:

$$\Delta E_{kjit} = \Delta X_{kjit}'\beta + \gamma \Delta S_{kjt} + \Delta W_{kjt}'\rho + \Delta Z_{kt}'\lambda + \Delta u_{kjit} \quad (3)$$

where $\Delta$ is the symbol for first differencing. For example in our data, $\Delta E_{kjit}$ stands for $E_{kji1994} - E_{kji1992}$ for panel 1.

When calculating the first difference, the constant term $\alpha_0$ and the fixed effect $c_{kji}$ are eliminated in equation 4.3. But the first-difference has an important weakness with unbalanced panels (Roodman, 2006a). Indeed, it subtracts the previous observation from the contemporaneous one which magnifies gaps in unbalanced panels. If a variable is missing at time $t$, then both this variable and its past realization are missing in the transformed data. Since there are only 2 waves in our data, first differencing our variables cost us one wave of data and subsequently decreases significantly the degree of freedom of the estimation.

$\Delta S_{kjt}$ could still be endogenous, that is $E(\Delta S_{kjt} \Delta u_{kjit}) \neq 0)$. The sources of this endogeneity must more likely be the same as with the social capital in level. The choice of instruments for $\Delta S_{kjt}$ is guided by two conditions: the instruments' lack of correlation with $\Delta u_{kjit}$ and their correlation with $\Delta S_{kjt}$. Note that $\Delta u_{kjit}$ in (4.3) contains information from two time periods in our data, $u_{kjit}$ and $u_{kjit-1}$ whereas $\Delta S_{kjt}$ contains $S_{kjt}$ and $S_{kjt-1}$.

We did not find a valid external instruments for social capital. So we draw instruments from within the data set. Natural candidate instrument is the social capital in levels with more than two lags as panel 1 that combines wave 1992 and wave 1994 has two years gap. Since the KHDS data is comprehensive, and the gap between wave is not constant, we were able to use the realization of social capital in 1991 to instrument for $\Delta S_{kjt}$.

When using the IV strategy, we estimate two-stage least squares (2SLS) regressions. The first stage regression is conducted through the first differenced variables with the following equation:

$$\Delta S_{kjt} = \Delta X_{kjit}'\beta + \phi S_{kjt-2} + \Delta W_{kjt}'\rho + \Delta Z_{kt}'\lambda + \Delta\mu_{kjit} \quad (4)$$

where $S_{kjt-2}$ is the value of social capital in $t-2$. Predicted $\hat{S}_{kjt}$ from this first-stage regression are then used for estimating equation (3) in the second stage. The value of social capital in 1991 appears to be valid instruments for $\Delta S_{kjt}$ as shown by the first stage result of our 2SLS estimation (see table 4.5).

# 4 Short term results

The econometric results for our short run analysis (panel 1) are presented in table 4.5. Column (1) gives the OLS results which we include for instructive concerns. Standard errors are made robust to heteroskedasticity and are adjusted for clustering at the community level. Considering both the results of the fixed effect estimator and that of the 2SLS regression, we find that social capital positively affects educational attainment in a higher magnitude than the OLS estimator.

## 4.1 Fixed effects results

Comparison of columns (1) vs. (2 ) of table 4.5 suggests that OLS estimates of social capital are strongly biased downwards, with the fixed effect point estimates greater than the OLS. OLS estimation not only biased the magnitude of the coefficients for many variables, it also altered the significance of the coefficients for some variables. For example, the sign for the variable measuring school quality is marginally positive and insignificant from the OLS regression but positive and significant from the fixed effect estimation. Thus, purging heterogeneity among individual units improve the results.

Table 5: Results 1992-1994

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| VARIABLES | OLS | FE | First st | 2SLS |
| Medium age | 1.27*** | 0.39*** | 0.48*** | 0.41* |
| | (0.075) | (0.083) | (0.147) | (0.237) |
| Old | 4.06*** | 0.98*** | 1.02*** | 0.93*** |
| | (0.095) | (0.116) | (0.250) | (0.326) |
| soc cap index | 0.12*** | 0.17*** | | 0.27* |
| | (0.033) | (0.028) | | (0.150) |
| At home | 0.93* | -0.39 | 0.98 | -0.45 |
| | (0.537) | (0.236) | (0.733) | (0.228) |
| Mother edu | 0.15* | -0.31 | 0.11 | 0.19 |
| | (0.082) | (0.130) | (0.281) | (0.189) |
| Log Expendt | 0.28*** | 0.41*** | 0.34*** | 0.34*** |
| | (0.077) | (0.072) | 0.111 | (0.104) |
| Hh size | 0.00 | 0.02 | 0.13** | 0.01 |
| | (0.009) | (0.022) | (0.056) | (0.034) |
| Numb teach | 0.01 | 0.12** | 0.15*** | 0.13* |
| | (0.017) | (0.046) | (0.044) | (0.070) |
| Soc cap 1991 | | | 0.21*** | |
| | | | (0.057) | |
| Constant | -2.78*** | -2.66*** | | |
| | (0.733) | (0.939) | | |
| Observations | 1,876 | 1,876 | 986 | 990 |
| R-squared | 0.505 | 0.339 | 0.203 | 0.307 |
| Number of pid91_94 | | 1,080 | | |
| sk11991 =0 | | | 13.69 | |

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.10

Fixed effect estimates presented in column (2) of table 4.5, based on children aged 7-16 years in panel 1, show a clear, significantly (at 1% level) a positive relationship between a child years of schooling attained and social capital. The estimate of social capital in column 2 implies

that increasing social capital index by one point would increase years of schooling by 0.17 year. Our findings are in line with previous researches that report positive effects of social capital on children education.

As with other studies,(Lioyd and Blanc, 1996; Glick and Sahn, 2000), we also find that household income proxied by per capita expenditure is significantly associated with children having more years of schooling.

The education level of parent, while having positive coefficient, is not significantly related to years of schooling. This seems counterintuitive, given the importance of a mother education in her child's educational outcome (see, e.g., Rosenzweig and Wolpin, 1994). We see two alternative explanations for this lack of significance. One may come from the fact that there exists a relatively higher number of missing values in parental education variables due among others to the relatively high rate of parental death and illiteracy in the Kagera sample. A second possibility can be attributed to the relatively low variability of this variable. This troubling finding between parental education and children' educational attainment is also supported in the literature among others by Feinstein and Symons (1999) and Abimbola (2007).

Child age was found to influence years of schooling. We consider distinct age ranges to determine whether there are structural differences between age groups. We include in the different regressions age of children (7-16 years old) as a three separate age groups represented by dummies for 7-9 (younger), 10-12 (Medium age) and 13-16 (old children). Most of the age dummies carries the right sign and are significants. The older age group (13-16) has the strongest effect. This suggests that as age of child increases, he tends to have more years of schooling. Young child may suffer from dropping out or late enrollment than older. For example, the withholding of young children from school in Tanzania has been shown to be a coping mechanism for families affected by the loss of an adult due to HIV/AIDS (Ainsworth *et al.*, 2005). Thus, any negative shock to household (income loss, parental loss and/or social capital loss) will likely to negatively impact young children enrollment, while older children (13-16) remain enrolled (Ainsworth *et al..*, 2005).

In contrast to what is usually found in the literature, the coefficient of household size and that of being at home while attending school are not statistically significants[10]. The quality of schooling measured by the number of teachers per school significantly affects school attainment. The fixed effect coefficient suggests that an additional teacher staffed in a Kagera school increases school attainment by 0.15 years. This result confirms that of Beegle and Ainsworth (2000).

[10]The empirical literature on the effects of family size on child education generally supports a negative relationship between family size and child education, even after controlling for socioeconomic factors. See for example Black *et al..*, 2005

## 4.2 2SLS results

As said before, the fixed effect results could still be biased because of the possible endogeneity of social capital index. As the availability of panel data provides an internal instrument, we explore a 2SLS strategy instrumenting social capital by its lagged value. Column (3) of Table 4.5 presents the first stage of the 2SLS estimation for our social capital index. The lagged value of social capital index in the first stage is correlated with the current social capital index and has the expected sign. The F-test shows that our instrument is valid as it passes the F-test threshold of 10 suggested by Staiger and Stock (1997).

Suspicion of the endogeneity of social capital is confirmed by 2SLS estimation results which shows a significant positive effects of social capital on child education. This 2SLS result reveals that the coefficient of the social capital index is larger than their OLS and fixed effect counterparts. It indicates that one point increases in social capital index would increase years of schooling by 0.27 year. This results strongly confirms the recent findings by Ainsworth (2002) that the amount and quality of social capital predict educational outcomes. Hence, following Ainsworth (2002), we are able to conclude that children who grow up in a rich social capital environment are more likely to be exposed to helpful social networks that may be educationally beneficial.

### 4.2.1 Subsamples

The impact of social capital in schooling outcomes in the short term may differ by socio-demographic group. We explore such differences by estimating the effect of social capital in education using different subsamples. The following table replicates the results. This exercise reveals that there are no crucial differences affecting the results discussed above and that our findings do not seem to be bound to a specific subsample. Hence, considering subsamples does not add any further predictive power in our model.

Columns (1-2) of table 4.6 show that the impact of social capital in educational attainment for boys and girls is broadly similar in the fixed effect regression, with the two subsamples having point estimates of 0.20 for the girls subsample and 0.19 for the boys subsample and both significants at 1% level.

But, the 2SLS results give another picture. Columns (5-6) of the same table show differences in the impact of social capital in the two subsamples, with the effect of social capital being stronger for girls (0.58 and significant at 5% level) than boys. This may suggest that social capital seems to advantage girls education in that if both boys and girls were in school, girls will have more years of schooling. T-tests point toward significant differences in mean schooling in

Table 6: Results 1992-1994/subsamples

| VARIABLES | (1) FE/Boys | (2) FE/Girls | (3) FE/Biol | (4) FE/foster | (5) 2SLS/Boy | (6) 2SLS/girls | (7) 2SLS/Biol | (8) 2SLS/fost |
|---|---|---|---|---|---|---|---|---|
| Medium age | 0.27 | 0.49*** | 0.28 | 0.21 | 0.46 | 0.40* | 0.31 | 0.43 |
| | (0.261) | (0.181) | (0.247) | (0.204) | (0.329) | (0.219) | (0.277) | (0.360) |
| Old | 0.89*** | 1.06*** | 0.91*** | 0.63** | 1.19*** | 0.75** | 0.93** | 0.73 |
| | (0.266) | (0.216) | (0.264) | (0.239) | (0.432) | (0.327) | (0.347) | (0.509) |
| At home | 0.01 | -0.54 | -0.06* | -0.63*** | -0.41** | -0.77** | -0.24 | -1.22*** |
| | (0.213) | (0.403) | (0.032) | (0.236) | (0.159) | (0.349) | (0.171) | (0.274) |
| Mother edu | -0.11 | 0.28 | 0.09 | -0.04 | -0.12 | 0.45* | 0.14 | 0.23 |
| | (0.180) | (0.223) | (0.207) | (0.301) | (0.215) | (0.241) | (0.236) | (0.341) |
| Soc cap index | 0.19*** | 0.20*** | 0.16*** | 0.32*** | 0.08 | 0.58** | 0.23 | 0.59** |
| | (0.042) | (0.042) | (0.037) | (0.058) | (0.233) | (0.219) | (0.212) | (0.274) |
| log Expendt | 0.35*** | 0.36*** | 0.49*** | 0.01 | 0.40*** | 0.28** | 0.45*** | 0.11 |
| | (0.106) | (0.111) | (0.092) | (0.129) | (0.130) | (0.135) | (0.123) | (0.150) |
| Hh size | -0.01 | 0.06** | 0.01 | 0.03 | 0.00 | 0.03 | 0.00 | 0.01 |
| | (0.031) | (0.029) | (0.035) | (0.028) | (0.047) | (0.034) | (0.050) | (0.030) |
| Number teach | 0.14* | 0.14** | 0.15** | 0.10 | 0.17** | 0.09 | 0.15* | 0.08 |
| | (0.072) | (0.062) | (0.068) | (0.065) | (0.082) | (0.077) | (0.075) | (0.089) |
| Constant | -2.08 | -2.90** | -3.66*** | 1.50 | | | | |
| | (1.563) | (1.364) | (1.246) | (1.914) | | | | |
| Observations | 786 | 726 | 1,040 | 584 | 517 | 483 | 657 | 343 |
| R-squared | 0.354 | 0.393 | 0.378 | 0.423 | 0.296 | 0.345 | 0.329 | 0.280 |

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.10

favor of girls than boys (3.3 compare to 2.7 years of schooling on average). This confirms earlier studies that argue that girls tend to get more schooling in Tanzania (see e.g Filmer, 2000).

The composition of household between biological and foster children offers another important heterogeneity issue in our sample. For example, in Sub-Saharan Africa, high levels of child fostering result in many children living in households of relatives than with their parents, even when both are alive (Ainsworth, 1996). The rate of fostering of children is high in Tanzania (32%). More often, fostering is linked to schooling decisions (Akresh, 2004a; Akresh, 2004b), children might be sent to live with relatives who are closer to schools or to have access to better quality schools (Gould, 1985). Child fostering also takes place in response to shocks, such as the death of one or both parents. At the time of the survey, Tanzania was strongly hit by the AIDS epidemic, with adult mortality rates three times higher than would normally be expected.

Our results indicate that biological children are not much privileged by the effect of social capital than foster children (with the fixed effect social capital coefficients 0.16 and 0.32 respectively, both significant at 1%). These results are confirmed in the 2SLS regression showing a larger and significant impact of social capital on the foster children group (see column 7-8 of table 4.6). This may suggest that, orphan and other disadvantageous children who could not have same education, may actually took advantage of the opportunity offered by the social capital present in their host households and gain some years of schooling. These findings confirm that of Ainsworth, Beegle and Koda (2002). Indeed, extended families and support networks may be effectively fostering children or transferring resources so that children can stay in school when a death or other shocks occur. Thus, these results are in straight line with the fulfillment of the African proverb "it takes a village to raise a child".

## 4.3 Robustness

We perform several robustness checks which results are presented in table 4.10 in Appendix B. The purpose of these analyzes is not only to prove the sensitivity of our results, but also to provide some evidence on the channels through which social capital may operate.

As discussed in the introduction, family's social capital may benefit educational outcomes through several mechanisms. On one hand, parents' connections to resourceful persons (relatives , friends and neighbors) can provide access to important resources such as information, money, control, opportunities, and potentially helpful contacts for children (Hoeffert, 1995; Answorth, 2002). On the other hand, family structure (measured here by the presence of two parents in the household) are expected to provide more time, more social control, more social support to children (Coleman ,1988; Teachman *et al.*, 1997; Sandefur *et al.*, 1999).

The data from the KHDS surveys do not allow the identification of the channels through which social capital affects education. In absence of these data, we adopt an indirect strategy by developing alternatives indexes of social capital. We expect that at least one of these alternative index will confirm our results.

Hence, we examine three alternatives indexes of social capital;

1. For comparison purposes, we relabel our initial index of social capital, "social capital index 1".

2. Social capital index 2 combines 3 indicators of social capital: family relationships with relatives, family relationships with friends/neighbors and remittances.

3. Social capital index 3 combines 3 indicators of social capital: family relationships with relatives, family relationships with friends/neighbors and indicator of family structure (presence of two parent in the household)

4. Social capital index 4 combines 2 indicators: family relationships with relatives and family relationships with friends/neighbors

Table 4.10 in appendix B reports the fixed effects results including one index of social capital by regression. OLS estimates (not reported) show a positive and significant relationships between all the social capital indexes and years of schooling. As with our preferred estimation (with social capital index 1), introducing fixed effects greatly improve the social capital effect for all the the indexes. The points estimates are more than the double of the corresponding OLS estimates. But the 2SLS results are different. Indeed, all social capital indexes have the expected signs, but only the estimates for the instrumented social capital index 4 built through family relationships and remittances ( see col 4 of table 4.10 in appendix B) have a strong and statistically significant coefficient as in our preferred estimation.

The finding that family relationships with relatives, friends and neighbors as well the receipt of remittances - as a measure of the strength of these relationships - significantly impact years of schooling leads us to conclude that information, help and assistance from neighbors may more likely be the channels through which social capital operates. This confirms our results that in Kagera, it effectively takes a village to raise a child. A contrario, family structure seems not to be one of the channel in the present context. When performing a regression analysis to assess the impact of each indicator of social capital separately, we find that that social capital plays a more influential role in education when considered as an index than as separate indicators (see table 4.12 in appendix B).

# 5 Persistence

The previous section proposes a contemporaneous relationship between social capital and children' educational outcome. However, social capital impact on education may not occur immediately but instead over future time periods.

The model to be estimated in this section empirically employs a panel data for two times periods spanning 10 years. Precisely, as stated earlier, we develop our long term empirical analysis on a two-wave panel dataset consisting of the last wave of KHDS I (completed in 1994) and wave 5 of KHDS II, completed in 2004.

The first-differenced estimation approach will be used to control for the individual unobservable heterogeneity bias. Thus, the question whether social capital has persistent effects on educational attainment over a long period of time can empirically be investigated by regressing the change of the depend variable on the change of the independents variables between 2004 and 1994, except the social capital index, which is used in level.

Indeed, the long term model includes the beginning-of period social capital index (we use the 1994 index of social capital) instead of the change of social capital. Performing the analysis in this way minimize the endogeneity biais, but this could still be a problem. Indeed, social capital in the beginning of the period may still be endogenous and this issue will be discussed later. All regressions have robust standard errors clustered at the community level. To be consistent with our short term analysis, the index of social capital is again evaluated at the family level. We exclude parents' education variables from the long term regression because of the high rate of missing values and also this variable does not change much between the two waves.

We introduce the following equations to explain how we address the issue of persistent effect of social capital[11]

$$\Delta E_{kjit} = \Delta X_{kjit}'\beta + \gamma S_{kjt-1} + \Delta W_{kjt}'\rho + \Delta Z_{kt}'\lambda + \Delta u_{kjit}$$

All the variables and parameters are as defined in the previous section and $\Delta E_{kjit}$ measures the change in educational attainment. The parameter $\gamma$ measures the effect of social capital in 1994 on the change of a child's education between 1994 and 2004.

There are familiar reasons that social capital is endogenous. Likewise, data on social capital capital are likely mis-measured. The classical solution for endogeneity is instrumental variables.

[11] We have also analyzed the impact of social capital in 1994 in children education in 2004. We include in this regression a set of individual, household, community control variables measured at 2004. Results are presented in appendix C, tables 4.14 and 4.15. The 2SLS results in column 3 of table 4.15 show that social capital in 1994 affect positively and significantly education in 2004.

Our long term analysis requires the change of the error term to be uncorrelated with the present and all past values of the social capital index. When sequential exogeneity holds, many instruments are available from within the system.

Indeed, we may include as instrument the realization of social capital for those time periods that are assumed to be unrelated to the differenced error term. Let us recall that sequential exogeneity means that $E[S_{kjt}u_{kjis}] = 0$ for $t \leq s$ and $\neq 0$ for $t > s$.

We use $S_{kjt-2}$, the level of social capital in 1993 to instrument $S_{kjt-1}$. Given the sequential exogeneity assumption, our first-differences equation , with instrumental variables will be unbiased if the instrument if $S_{kjt-2}$ is unrelated to $\Delta\mu_{kjit}$ for $t > s$. More explicitly, $S_{kjt-2}$ must not be correlated with future $\mu_{kjit}$ or $\mu_{kjit-1}$. Considering the results of the first stage regression in table 4.7 $S_{kjt-2}$ appears to be a valid instrument for $S_{kjt-1}$ in panel 2.

## 5.1 Long term results

We report in table 4.7 the results for the first differences OLS model (FD-OLS) and the first differences 2SLS model (FD-2SLS).

The 2SLS estimates indicate that social capital in 1994 has a meaningful and strong effect on education change between 1994 and 2004 as shown in column 3 of the table 4.7. These estimates indicate that social capital in 1994 increases the change in education by some 1.12 year in a period of 10 years. Comparing the FD-OLS and FD-2SLS coefficients of social capital , shows that instrumentation strengthens the positive and significant social capital effect on educational change.

Likewise, the change of education between 2004 and 1994 is also a function of changes in a series of the other control variables included in the regression. However, among these variables, we find that the change in the number of teachers in Tanzanian school does not affect change in education. That may indicate that there is a severe shortage of teachers in Tanzanian schools as argued earlier.

The seminal work of Coleman (1988) raised the hypothesis that social capital contributes to human capital formation. There is however less evidence in the literature on the long-term impact of social capital on education because there are very few longitudinal studies in developing countries. As our data covers a sufficiently long period, we were able to observe the longer-term effects of social capital on education. Hence, controlling for all else, there is evidence of lasting effects of social capital on children education ten years later.

There is no substantial variation in the magnitude of the coefficient using different subgroups with the coefficient of social capital estimate around 1.10 in the different subsamples. Interest-

Table 7: result 1994-2004

| | (1) | (2) | (3) |
|---|---|---|---|
| VARIABLES | Ols FD | First Stage FD | 2SLS FD |
| D.Middle age | 3.58*** | -0.02 | 3.91*** |
| | (0.510) | (0.186) | (0.494) |
| D. Old | 6.19*** | -0.02 | 6.66*** |
| | (0.793) | (0.248) | (0.829) |
| D. Athome | 0.33** | 0.08 | 0.35** |
| | (0.148) | (0.056) | (0.157) |
| D. Expend | 1.41*** | -0.06* | 1.43*** |
| | (0.180) | (0.031) | (0.197) |
| D. Hh size | -0.05 | -0.07* | -0.03 |
| | (0.104) | (0.033) | (0.106) |
| D. Numb teach | -0.00 | 0.01 | -0.00 |
| | (0.056) | (0.008) | (0.059) |
| Soc cap 1994 | 0.44** | | 1.12** |
| | (0.195) | | (0.467) |
| Soc cap 1993 | | 0.46*** | (0.467) |
| | | (0.058) | |
| Observations | 1,245 | 1,126 | 1,114 |
| R-squared | 0.322 | 0.237 | 0.333 |
| sk11993 =0 | | 63.60 | |

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.10

ingly, splitting the results by gender for instance, reveals that the relationship between social capital and education in the long term is not driven by the sex of the child, contrary to what we find in our short term regression. Thus, social capital appears as an important factor in correcting gender bias in education (see Atemnkeng 2005; 2010) over the long term (see table 4.13 in appendix C).

Table 8: 2SLS 1994-2004 by soc cap indicators

| | (1) | (2) |
|---|---|---|
| VARIABLES | 2SLS FD3 | 2SLS FD4 |
| D.Middle age | 2.02*** | 1.69** |
| | (0.543) | (0.669) |
| D. Old | 3.04*** | 2.64*** |
| | (0.718) | (0.804) |
| D. Athome | 0.16 | -0.31* |
| | (0.159) | (0.177) |
| D. Expend | 0.34** | -0.33* |
| | (0.155) | (0.199) |
| D. Hh size | -0.34** | -0.25* |
| | (0.146) | (0.131) |
| D. Numb teach | -0.01 | -0.04 |
| | (0.031) | (0.058) |
| D. log remit 1994 | | 1.17*** |
| | | (0.109) |
| D. Relativs 1994 | 0.61*** | |
| | (0.082) | |
| Observations | 1,086 | 1,114 |
| Relative 1993 =0 | 13.69 | |
| Remittances 1993 =0 | | 119.45 |

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.10

Relationships with relative and receiving remittances seem to be the mechanisms that sustain the long term effect of social capital in education, with both indicators being significants at 1 percent level (see table 4.8) . The 2SLS coefficients of the two remaining indicators of social capital (two parents families and Friends' network) are insignificants (results not reported here).

# 6 Conclusion

This paper examines the role of social capital on the educational attainment. According to Coleman (1988), parent human capital and financial capital are made available to children through their family social capital.

We conceptualize social capital at the family level and, exploiting data from the Kagera region in Tanzania, we analyze the short and long term impacts of family social capital on educational attainment in Tanzania. This study differs from previous ones in that it accounts also for the long term effects of social capital in education.

We were able to identify four proxies of social capital in the KHDS data: household relative's network, household friends and neighbors network, the amount of remittances received and whether both parents are present at home. These indicators were subjected to a Principal Components Analysis to create an index of social capital.

We observe that in the short-run as well in the long term, households endogenously determine their level of social capital. Thus, we used instrumental variables to account for this endogeneity and to determine the causal effects of social capital in education.

Our short run empirical results confirm and extend in developing countries earlier findings suggesting that social capital affects educational outcomes of children. As our data covers a sufficiently long period, we were able to observe the longer-term effects of social capital on education. Hence, controlling for all else, there is evidence of lasting effects of social capital on children education ten years later.

Analyzing our results, we contend that information, help and assistance from friends and neighbors may more likely be the channels through which social capital impacts child educational outcomes. This confirms the African proverb that in Kagera, it effectively "takes a village to raise a child". Further, social capital plays a more influential role in education when considered as an index than as separate indicator.

Although this study extends the findings of previous research in developing countries, it is not without certain limitations about the measurement of social capital. The indicators used in this study managed to capture only a part of the total social capital that may be available to the student. There is rooms for improvement in this area.

However, the important contribution of social capital to student educational attainment discussed and empirically tested in this study is expected to remain at the center of future research.

# References


[1] Abimbola, T., (2007), *Malaria, labor supply, and schooling in Sub-Sahara Africa.* unpublished paper, University of South Florida, .

[2] Ainsworth, M., Beegle K., and Koda G., (2002). *The Impact of Adult Mortality on Primary School Enrollment in Northwestern Tanzania.* African Region Human Development Working Paper Series, No. 23961. Washington D.C.: The World Bank.

[3] Akresh, R. (2004), *Adjusting household structure: school enrollment impacts of child fostering in Burkina Faso.* Discussion Paper no. 897, Economic Growth Center, Yale University.

[4] Alderman, H., J.R. Berman, H.-P. Kohler, J.A. Maluccio, and S.C. Watkins., (2001). *Attrition in Longitudinal Household Survey Data: Some Tests for Three Developing-Country Samples.* Demographic Research 5(4).

[5] Al-Samarrai, S. and Peasgood T., (1998). *Education Attainments and Household Characteristics in Tanzania.* Economics of Education Review 17(4): 395-417.

[6] Altonji, J. G., Hayashi, F., and Kotlikoff, L. J., (1992). *Is the Extended Family Altruistically Linked?* Direct Tests Using Micro Data. The American Economic Review, 82(5):11771198.

[7] Anderson, J.B. (2008). *Social capital and student learning: Empirical results from Latin American primary schools.* Economics of Education Review, 27, 439449.

[8] Asadullah, M. N., (2008), *Sense in sociability? Maternal education, social capital and child schooling in rural Bangladesh.* European Journal of Development Research

[9] Bassani, C., (2006). *A test of social capital theory outside of the American context: Family and school social capital and youths math scores in Canada, Japan, and the United States.* International Journal of Educational Research, 45(6), 380-403.

[10] Beegle, K., (1999). *The Impact of Adult Mortality in Households: Evidence from Rural Tanzania.* Mimeo. Santa Monica, CA: RAND Corporation.

[11] Beegle, K., De Weerdt J., and Dercon S., (2006a *Kagera Health and Development Survey (KHDS), 2004, Basic Information Document.* Document. World Bank, Washington, DC.

[12] Beegle, K., De Weerdt J., and Dercon S., (2006b). *Orphanhood and the Long-Run Impact on Children.* American Journal of Agricultural Economics 88:126672.

[13] Beegle, K., de Weerdt J., and Dercon S., (2010). *Orphanhood and Human Capital Destruction: is there persistence into adulthood.* Demography, 47(1).

[14] Bommier, A., Lambert, S., (2000). *Education demand and age at school enrollment in Tanzania.* Journal of Human Resources XXXV, 177 203.

[15] Buchel, F., and Duncan G., (1998). *Do Parents Social Activities Promote Childrens School Attainments? Evidence from the German Socioeconomic Panel.* Journal of Marriage and the Family 60(1)

[16] Burke, K., Beegle, K., (2004). *Why children arent attending school: the case of rural Tanzania.* Journal of African Economies 13, 333 355.

[17] Cameron, C.A,. Colin and Trivedi P. K., (2005), *Microeconometrics : Methods and Applications,* Cambridge University Press, Cambridge.

[18] Cameron, C.A., and Trivedi P. K., (2010), *Microeconometrics Using Stata,* Stata Press, College Station, TX.

[19] Cogneau, D., and Jedwab, R. (2008). *Family Income and Child Outcomes: The 1990 Cocoa Price Shock in Cote d'Ivoire.* Dveloppement, Institutions and Analyses de Long terme.

[20] Coleman, J. (1988), *Social capital in the creation of human capital*, American Journal of Sociology 94 (Suppl.), S95-S120.

[21] Cox, D., and Fafchamps M.,(2007), *Extended Family and Kinship Networks: Economic Insights and Evolutionary Directions, in Handbook of Development Economics.* (Eds.T.P. Schultz and R.E. Evenson). Volume 4, Amsterdam: Elsevier Press.

[22] Deininger K., (2001). *Does cost of schooling affect enrollment by the poor? Universal primary education in Uganda.* Washington, DC: World Bank.

[23] Duru-Bellat, M., (2004), *L'cole des filles : quelle formation pour quels rles sociaux ?* Paris, L'Harmattan.

[24] Feinstein, L. and Symons, J. (1999). *Attainment in secondary school. Oxford Economic Papers*, 51, 300-321.

[25] Filmer, Deon. 2000. *The Structure of Social Disparities in Education: Gender and Wealth.* Policy Research Working Paper No. 2268, World Bank Development Research

Group/Poverty Reduction and Economic Management Network. Washington, D.C.: World Bank.

[26] Glewwe, P, and Hanan J., (1994), *Student Achievement and Schooling Choice in Low-Income Countries: Evidence from Ghana* Journal of Human Resources, 29(3): 843-864.

[27] Glewwe, P. (2002). *Schools and Skills in Developing Countries: Education Policies and Socioeconomic Outcomes.* Journal of Economic Literature 40: 436-482.

[28] Glick, Peter and David E. Sahn, 2000, *Schooling of girls and boys in a West African Country: the effects of parental education, income and household structure.* Economics of Education Review, 19, pp.63-87.

[29] Greene, William H. (2007). *Econometric Analysis,* 6th Edition. New Jersey: Prentice Hall.

[30] Hagen, Jens, Mahmoud, Toman Omar, and Trofimenko, Natalia (2010).*Orphanhood and Critical Periods in Childrens Human Capital Formation: Long-Run Evidence from North-Western Tanzania.* Kiel Institute for the World Economy, September, No. 1649.

[31] Hanushek, E., (1995). *Interpreting recent research on schooling in developing countries,* World Bank Rechearch Observer.

[32] Hofferth, S. L., Boisjoly, J., and Duncan, G. J. (1998). *Parents extrafamilial resources and childrens school attainment.* Sociology of Education, 71, 246268.

[33] Israel, G. D., Beaulieu, L. J., and Hartless, G. (2001). *The influence of family and community social capital on educational achievement.* Rural Sociology, 66, 43-68.

[34] Jger, Mads, and Anders Holm, (2007). *Does parents economic, cultural, and social capital explain the social class effect on educational attainment in the Scandinavian mobility regime?* Social Science Research 36:719-744.

[35] KHDS (2004). *Users Guide to the Kagera Health and Development Survey Datasets.* Development Research Group. The World Bank.

[36] Lauer, C.,(2003). *Family background, cohort and education: A French-German comparison based on a multivariate ordered probit model of educational attainment.* Labour Economics, 10(2): 231-251

[37] Lloyd, C. B, and BLANC K. (1996). *Children's Schooling in sub-Saharan Africa: The Role of Fathers, Mothers, and Others.*Population and development review,Vol. 22, N 2, p. 265-298.

[38] Morgan, S., and Sorensen, A. (1999). *Parental networks, social closure, and mathematics learning: A test of colemans social capital explanation of school effects.* American Sociological Review, 64(5), 661682.

[39] Nichols, A., (2007). *Causal inference with observational data.*, Stata Journal 7 (4),507541.

[40] Mouw T., (2006). *Estimating the causal effect of social capital: a review of recent research.* Annual Review of Sociology 32: 79102

[41] Roodman, D. (2006a). *How to Do xtabond2: An Introduction to Difference and System GMM in Stata,.*Center for Global Development Working Paper Number 103.

[42] Rosenzweig, M.R., Wolpin, K.I., 1993. *Credit market constraints, consumption smoothing, and the accumulation of durable production assets in low-income countries: investments in Bullocks in India.* Journal of Political Economy 101 (2), 223244, April.

[43] Runyan, D.K., and C. Gould. (1985) *Foster Care for Child Maltreatment: Impact on Delinquent Behavior.* Pediatrics. 75: 562-68.

[44] Sandefur, G., Meier, A., and Hernandez, P. (1999). *Families, social capital and educational continuation.* CDE Working Paper No. 99-19: Center for Demography and Ecology. Madison, WI: University of Wisconsin-Madison.

[45] Sandefur, Gary D., Meier A., .Campbell M. (2006). *Family Resources, Social Capital, and College Attendance.* Social Science Research 35:52553.

[46] Simonsen, M., and Skipper, L., (2007) *Child Health in a Developing Country: Consequences for Short- and Medium Term Outcomes.* Mimeo, University of Arhus and Institute for Local Government Studies.

[47] Stanton-Salazar, R., and Dornbusch, S.M., (1995). *Social capital and the reproduction of inequality: Information networks among Mexican-origin high school students.* Sociology of Education, 68, 116-135

[48] Suryadarma, D., Suryahadi A., Sudarno S., and F. Halsey R., (2006). *Improving Student Performance in Public Primary Schools in Developing Countries: Evidence from Indonesia.* Education Economics, 14(4): 401-429.

[49] Teachman J. D., Paasch K., and Carver K., (1997), *Social Capital and the Generation of Human Capital*, Social Forces, 75, 1343-1359.

[50] Wedgwood, R. (2005). *Education and poverty reduction in Tanzania,* in 8th Oxford International Conference on Education and Development, 13-15 September, New College: Oxford. Available at http://www.cas.ed.ac.uk/research/pbet.html.

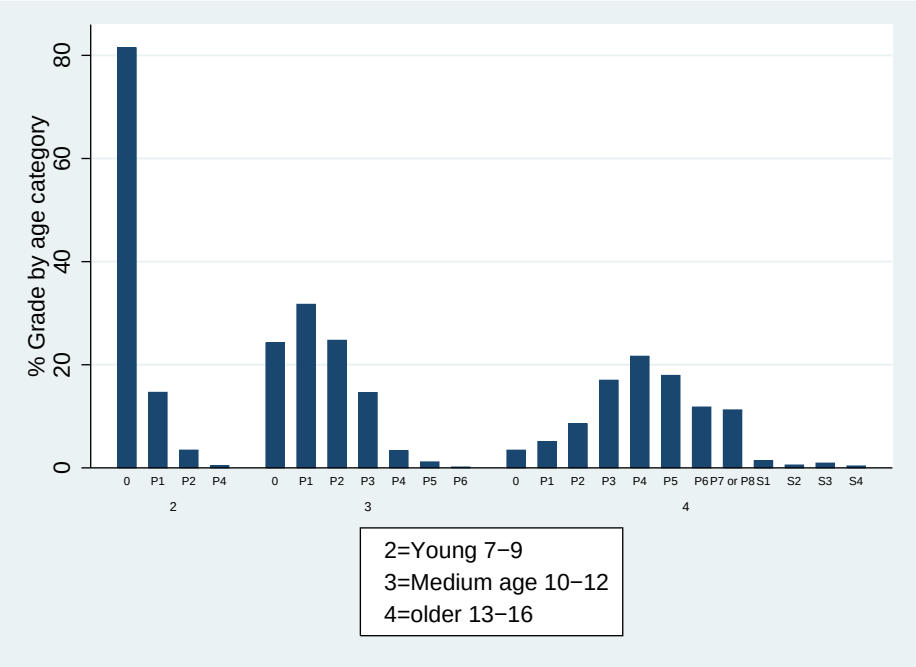
[51] Wooldridge, Jeffrey M., (2002). *Econometric Analysis of Cross Section and Panel Data.* Cambridge, Massachusetts: MIT Press.

# Appendix A: Variables definition

Table 9: Variables definitions

| Variables | Definition |
|---|---|
| Yrsed | Years of education |
| Ageyr | Age in years |
| Grpage | Age category: Young: 7-9 ans; Medium age (Med): 10-12ans; Old: 13-16 ans |
| Sk1 | Social capital index |
| Athome | At home while attending school |
| Schldist | Distance to School |
| Mother educ | Dummy variable indicating if mother has some education or not |
| log expendit | Log per capita expenditure |
| hhsize | Household size |
| Nblack | Number of blackboards per class |
| Nteachers | Number of teachers |
| Attended | Whether attended school or not (1994-2004) |
| sk1hats | OLS predicted values of social capital |
| D.variables | Variable in first difference |
| P | Primary school grade |

Figure 2: Highest grade by age group 1992-1994



# Appendix B: Robustness checks

Table 10: Fixed effects 1992-1994 Alternative social capital indexes

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| VARIABLES | FE | FE | FE | FE |
| Medium age | 0.34* | 0.17 | 0.17 | 0.29 |
| | (0.179) | (0.172) | (0.172) | (0.176) |
| Old | 0.95*** | 0.72*** | 0.71*** | 0.89*** |
| | (0.188) | (0.182) | (0.182) | (0.185) |
| Athome | -0.35 | -0.17 | -0.17 | -0.32 |
| | (0.287) | (0.266) | (0.267) | (0.287) |
| Mother edu | -0.04 | 0.03 | 0.03 | -0.02 |
| | (0.157) | (0.152) | (0.152) | (0.156) |
| log Expendt | 0.31*** | 0.30*** | 0.30*** | 0.31*** |
| | (0.094) | (0.088) | (0.088) | (0.093) |
| Hh size | 0.02 | -0.00 | -0.00 | 0.01 |
| | (0.026) | (0.026) | (0.027) | (0.026) |
| Number teach | 0.14** | 0.13** | 0.13** | 0.14** |
| | (0.060) | (0.060) | (0.060) | (0.060) |
| Social cap index 1 | 0.21*** | | | |
| | (0.043) | | | |
| Social cap index 2 | | 0.23*** | | |
| | | (0.039) | | |
| Social cap index 3 | | | 0.24*** | |
| | | | (0.041) | |
| Social cap index 4 | | | | 0.19*** |
| | | | | (0.036) |
| Constant | -2.07 | -1.55 | -1.53 | -1.93 |
| | (1.263) | (1.221) | (1.234) | (1.255) |
| Observations | 1,278 | 1,440 | 1,279 | 1,439 |
| R-squared | 0.364 | 0.394 | 0.395 | 0.373 |
| Number of pid91_94 | 752 | 913 | 752 | 913 |

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.10

Table 11: 2SLS 1992-1994 Alternative social capital indexes

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| VARIABLES | 2SLSsk1 | 2SLSsk3 | 2SLSk4 | 2SLSk5 |
| D.Medium age | 0.35 | 0.52* | 0.50* | 0.32 |
| | (0.243) | (0.280) | (0.262) | (0.244) |
| D.Old | 0.86** | 1.21*** | 1.19*** | 0.80** |
| | (0.332) | (0.428) | (0.392) | (0.332) |
| D.Athome | -0.52** | -0.26 | -0.25 | -0.48* |
| | (0.256) | (0.220) | (0.219) | (0.238) |
| D.log Expendt | 0.29** | 0.41*** | 0.41*** | 0.28** |
| | (0.111) | (0.125) | (0.118) | (0.108) |
| D.Hh size | 0.00 | 0.03 | 0.03 | -0.00 |
| | (0.034) | (0.046) | (0.040) | (0.035) |
| D.Number teach | 0.13* | 0.19** | 0.19*** | 0.12* |
| | (0.070) | (0.076) | (0.058) | (0.070) |
| Soc IV1 | 0.34* | | | |
| | (0.184) | | | |
| Soc IV2 | | 0.04 | | |
| | | (0.220) | | |
| Soc IV3 | | | 0.05 | |
| | | | (0.189) | |
| Soc IV4 | | | | 0.31** |
| | | | | (0.146) |
| Observations | 776 | 806 | 806 | 797 |
| R-squared | 0.319 | 0.312 | 0.312 | 0.321 |

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.10

Table 12: Fixed effects 1992-1994 social capital indicators

| | (1) | (2) | (3) |
|---|---|---|---|
| VARIABLES | FE | FE | FE |
| Medium age | 0.12 | 0.15 | 0.53** |
| | (0.171) | (0.167) | (0.206) |
| Old | 0.66*** | 0.69*** | 1.23*** |
| | (0.181) | (0.175) | (0.227) |
| Relatives | 0.08*** | 0.08*** | |
| | (0.019) | (0.018) | |
| Friends/ Neighbors | 0.07*** | 0.07*** | |
| | (0.019) | (0.018) | |
| Log remittances | -0.02 | | 0.04** |
| | (0.016) | | (0.015) |
| Athome | -0.15 | -0.20 | -0.38 |
| | (0.276) | (0.280) | (0.269) |
| Dist to school | -0.00 | -0.00 | -0.00 |
| | (0.001) | (0.001) | (0.001) |
| Mother edu | 0.05 | 0.04 | -0.11 |
| | (0.148) | (0.151) | (0.161) |
| log Expendt | 0.28*** | 0.28*** | 0.39*** |
| | (0.092) | (0.092) | (0.092) |
| Hh size | -0.00 | -0.00 | 0.03 |
| | (0.027) | (0.027) | (0.026) |
| number teach | 0.13** | 0.13** | 0.18*** |
| | (0.061) | (0.060) | (0.061) |
| Constant | -1.46 | -1.57 | -3.75*** |
| | (1.310) | (1.302) | (1.070) |
| Observations | 1,278 | 1,440 | 1,439 |
| R-squared | 0.402 | 0.398 | 0.313 |
| Number of pid91_94 | 752 | 913 | 913 |

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.10

# Appendix C: Long term results 1994-2004

Table 13: 2SLS/Sub-samples

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| VARIABLES | 2SLS/FD/Boys | 2SLS/FD/Girls | 2SLS/FD/Biol | (2SLS/FD/fost) |
| D.Middle age | 2.47*** | 5.99*** | 4.15*** | 3.82*** |
| | (0.536) | (0.708) | (0.597) | (0.773) |
| D.Old | 5.31*** | 8.65*** | 7.71*** | 6.08*** |
| | (0.782) | (1.211) | (1.314) | (1.026) |
| D. At home | 0.19 | 0.56** | -0.65** | 1.21*** |
| | (0.168) | (0.230) | (0.306) | (0.270) |
| D. Expendit | 1.51*** | 1.34*** | 1.40*** | 1.31*** |
| | (0.171) | (0.283) | (0.213) | (0.288) |
| D.Hh size | -0.03 | -0.02 | -0.16 | 0.03 |
| | (0.102) | (0.130) | (0.141) | (0.127) |
| D.Number teach | 0.02 | -0.02 | 0.01 | -0.02 |
| | (0.057) | (0.064) | (0.076) | (0.045) |
| Soc cap 1994 | 1.12** | 1.19** | 1.20** | 1.09** |
| | (0.500) | (0.537) | (0.586) | (0.426) |
| Observations | 584 | 530 | 506 | 608 |
| R-squared | 0.355 | 0.329 | 0.329 | 0.391 |

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.10

Table 14: Fixed effects by sc indicators

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| VARIABLES | Ols FD | Ols FD | Ols FD | Ols FD |
| D. Medium age | 2.84*** | 3.52*** | 3.40*** | 2.84*** |
| | (0.491) | (0.530) | (0.551) | (0.483) |
| D.Old | 4.86*** | 6.10*** | 5.54*** | 4.68*** |
| | (0.642) | (0.805) | (0.790) | (0.624) |
| D. Athome | 0.13 | 0.61*** | 0.51*** | 0.26** |
| | (0.114) | (0.184) | (0.130) | (0.127) |
| D. log expendit | 0.81*** | 1.39*** | 1.18*** | 0.90*** |
| | (0.157) | (0.175) | (0.197) | (0.168) |
| D. Hh size | -0.14 | -0.10 | -0.16 | -0.22** |
| | (0.085) | (0.111) | (0.114) | (0.081) |
| D. Number teach | -0.02 | 0.00 | 0.00 | -0.00 |
| | (0.030) | (0.054) | (0.054) | (0.028) |
| Log remittances 1994 | 0.40*** | | | |
| | (0.050) | | | |
| Two parents 1994 | | 0.70** | | |
| | | (0.264) | | |
| Friends/neigh 1994 | | | 0.42*** | |
| | | | (0.089) | |
| Relative 1994 | | | | 0.30*** |
| | | | | (0.046) |
| Observations | 1,245 | 1,245 | 1,153 | 1,204 |
| R-squared | 0.478 | 0.312 | 0.384 | 0.501 |

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.10

Table 15: Soc cap 1994 vs variables 2004

| | (1) | (2) | (3) |
|---|---|---|---|
| VARIABLES | Ols 04-94 | FS 04-94 | 2SLS 04-94 |
| Middle age | 2.98*** | 0.03 | 2.82*** |
| | (0.226) | (0.121) | (0.205) |
| Old | 3.01*** | 0.09 | 2.93*** |
| | (0.303) | (0.154) | (0.270) |
| Soc cap 1994 | 0.16** | | 0.42** |
| | (0.066) | | (0.159) |
| At home2 | 0.13 | 0.02 | -0.05 |
| | (0.273) | (0.129) | (0.249) |
| Mother edu | 0.06 | -0.06 | 0.14 |
| | (0.184) | (0.139) | (0.178) |
| Log expendit | 0.02 | -0.02 | 0.02 |
| | (0.064) | (0.049) | (0.061) |
| HH size | 0.00 | -0.02 | -0.01 |
| | (0.028) | (0.016) | (0.027) |
| Number teach | 0.02 | 0.01 | 0.03 |
| | (0.018) | (0.016) | (0.022) |
| Soc cap 1993 | | 0.48*** | |
| | | (0.052) | |
| Constant | 3.30*** | 0.10 | 3.54*** |
| | (0.831) | (0.687) | (0.816) |
| Observations | 1,293 | 1,170 | 1,263 |
| R-squared | 0.331 | 0.219 | 0.285 |
| sk11993 =0 | | 85.08 | |

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.10

# Appendix D: Attritions and correlations tables

Table 16: Attrition 1992-1994

| Sample | variable | mean | min | max | N | t-test |
|---|---|---|---|---|---|---|
| Attriers | Age | 12.500 | 7.000 | 16.000 | 212.000 | 2.0245 (0.04) |
| | Years edu | 3.485 | 1.000 | 9.000 | 165.000 | 1.9514 (0.05) |
| | Grade | 2.712 | 0.000 | 9.000 | 212.000 | 2.7016 (0.01) |
| | Grade_head | 5.475 | 0.000 | 19.000 | 179.000 | 2.0915 (0.04) |
| | Father_edu | 5.742 | 0.000 | 19.000 | 93.000 | 1.5716 (0.12) |
| | Mother_edu | 4.815 | 0.000 | 11.000 | 81.000 | 2.0784 (0.04) |
| | Age_head | 51.075 | 14.000 | 95.000 | 212.000 | -0.3943 (0.69) |
| Non attr | Age | 12.123 | 7.000 | 16.000 | 822.000 | |
| | Years edu | 3.190 | 1.000 | 9.000 | 588.000 | |
| | Grade | 2.285 | 0.000 | 9.000 | 821.000 | |
| | Grade_head | 4.988 | 0.000 | 17.000 | 736.000 | |
| | Father_edu | 5.260 | 0.000 | 11.000 | 454.000 | |
| | Mother_edu | 4.164 | 0.000 | 11.000 | 427.000 | |
| | Age_head | 51.532 | 6.000 | 89.000 | 822.000 | |

Table 17: Attrition 1994-2004

| **Sample** | **variable** | **mean** | **min** | **max** | **N** | **t-test** |
|---|---|---|---|---|---|---|
| Attriers | Age | 9.625 | 0.000 | 29.000 | 352.000 | -7.8588 (0.00) |
| | Grade | 2.333 | 0.000 | 11.000 | 351.000 | -4.7253 (0.00) |
| | Grade_head | 5.608 | 0.000 | 17.000 | 181.000 | 2.5705 (0.01) |
| | Father_edu | 6.190 | 2.000 | 11.000 | 21.000 | 1.5558 (0.12) |
| | Mother_edu | 4.970 | 0.000 | 7.000 | 33.0000 | .3971 (0.69) |
| | Age_head | 47.552 | 17.000 | 98.000 | 212.000 | -1.2700 (0.20) |
| Non attr | Age | 12.789 | 0.000 | 30.000 | 3066.000 | |
| | Grade | 3.179 | 0.000 | 18.000 | 3053.000 | |
| | Grade_head | 5.064 | 0.000 | 17.000 | 1527.000 | |
| | Father_edu | 5.334 | 0.000 | 18.000 | 883.000 | |
| | Mother_edu | 4.775 | 0.000 | 11.000 | 839.000 | |
| | Age_head | 48.941 | 17.000 | 93.000 | 1772.000 | |

Table 18: OLS on Attrition 1994-2004

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| VARIABLES | Ols attr1 | Ols attr1 | Ols attr1 | Ols attr1 | Ols attr1 |
| Age | -0.03** | -0.03** | -0.03** | -0.03** | -0.03** |
| | (0.012) | (0.012) | (0.012) | (0.012) | (0.012) |
| Grade | 0.03 | 0.03 | 0.03* | 0.03* | 0.03 |
| | (0.018) | (0.017) | (0.017) | (0.017) | (0.018) |
| At home | -0.13 | -0.13 | -0.14 | -0.14 | -0.13 |
| | (0.098) | (0.099) | (0.099) | (0.098) | (0.098) |
| Tribe | -0.01 | -0.01 | -0.02 | -0.01 | -0.01 |
| | (0.015) | (0.014) | (0.015) | (0.015) | (0.015) |
| hh size | 0.02** | 0.02** | 0.02*** | 0.03*** | 0.02*** |
| | (0.008) | (0.008) | (0.008) | (0.008) | (0.008) |
| Social capital | 0.03 | | | | |
| | (0.022) | | | | |
| Relatives | | 0.02** | | | |
| | | (0.008) | | | |
| Friends | | | -0.00 | | |
| | | | (0.011) | | |
| Two parents family | | | | -0.05 | |
| | | | | (0.070) | |
| Remit | | | | | 0.01 |
| | | | | | (0.011) |
| Constant | 0.71*** | 0.69*** | 0.71*** | 0.70*** | 0.65*** |
| | (0.148) | (0.152) | (0.149) | (0.149) | (0.178) |
| Observations | 374 | 374 | 374 | 374 | 374 |
| R-squared | 0.049 | 0.055 | 0.041 | 0.043 | 0.044 |

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.10

Table 19: Cross-correlation social capital indicators 1994

| Variables | Remittances | Relalives | Friends and Neigh | Two Parents |
|---|---|---|---|---|
| Remittances | 1.000 | | | |
| Relalives | 0.617 | 1.000 | | |
| Friends and Neigh | 0.297 | 0.132 | 1.000 | |
| Two Parents | -0.104 | -0.094 | 0.016 | 1.000 |

Table 20: Cross-correlation social capital indicators 2004

| Variables | Remittances | Relalives | Friends and Neigh | Two Parents |
|---|---|---|---|---|
| Remittances | 1.000 | | | |
| Relalives | 0.219 | 1.000 | | |
| Friends and Neigh | 0.245 | 0.189 | 1.000 | |
| Two Parents | -0.046 | 0.142 | 0.056 | 1.000 |

Institut de Recherches Économiques et Sociales
Université catholique de Louvain

Place Montesquieu, 3
1348 Louvain-la-Neuve, Belgique


ISSN 1379-244X D/2013/3082/03