UNIVERSITY OF GOTHENBURG
SCHOOL OF BUSINESS, ECONOMICS AND LAW

WORKING PAPERS IN ECONOMICS

No 472 (revised)

The Push Towards UPE and the Determinants of the Demand for Education in Tanzania

Måns Nerman

Trudy Owens

March 2012
(First edition: October 2010)

ISSN 1403-2473 (print)
ISSN 1403-2465 (online)

Department of Economics
School of Business, Economics and Law at University of Gothenburg
Vasagatan 1, PO Box 640, SE 405 30 Göteborg, Sweden
+46 31 786 0000, +46 31 786 1326 (fax)
www.handels.gu.se info@handels.gu.se

The Push Towards UPE and the Determinants of the Demand for Education in Tanzania[1]

Måns Nerman[2]

Department of Economics, University of Gothenburg, Gothenburg, Sweden

Trudy Owens

Faculty of Social Sciences, University of Nottingham, Nottingham, United Kingdom

Abstract

This paper uses household data to investigate the determinants of demand for education in Tanzania and test whether these have changed during the government's push for Universal Primary Education in the 2000s. We find that the abolition of school fees was followed by an overall increase in enrolment, yet the sustained importance of household's consumption, livelihood and education indicates that the socio-economic standing of the household remains an important source of educational inequality. We also include estimated returns to education as an explanatory factor but find no indications that returns determine demand in Tanzania.

Keywords: primary education, household behaviour, Tanzania

JEL classification: I21, O15

[1] We are grateful to Niklas Bengtsson, Arne Bigsten, Måns Söderbom, and the participants at seminars at the Nordic Conference in development Economics in Helsinki and at the University of Gothenburg for useful comments.

[2] Corresponding author. *Postal Address*: P.O. Box 640, SE 405 30, Gothenburg, Sweden. *E-mail*: mans.nerman@economics.gu.se

1 Introduction

Achieving Universal Primary Education (UPE) is explicitly stated as one of the Millennium Development Goals and has been the focus of many policy makers in developing countries during the last decade. The benefits of increased education are well documented, not only at the individual level where education may provide a pathway out of poverty or improved health, but there is also evidence regarding social benefits of education such as higher growth levels and more rapid technology diffusion (see Rosenzweig (2010) for a recent discussion). Like many other developing countries, Tanzania has made a push towards UPE since the turn of the century and has seen both enrolment rates increase and attainment levels rise. We examine how the demand for education in Tanzania has developed during the recent UPE policy program. Drawing on existing theoretical and empirical literature, we aim to investigate if and how the importance of commonly suggested determinants of demand for education have changed in the new millennium. To this end we include variables to capture the costs of education, the benefits of education measured as the observed financial return to education in the economic context of the household, and, finally, the preferences for education.

Using data from two nationally representative household budget surveys covering mainland Tanzania in 2001 and 2007, we find evidence that both direct and opportunity costs are important determinants of educational demand, as is the household's level of consumption. In line with previous research we find that parents' education and the child's relationship to the head of household are important, which is indicative of the importance of household preferences. A key finding of the paper is the potential importance of social norms in determining demand. We find that the average level of education within the local community is a significant predictor of children's education, which indicates that educational choices are affected by the views on education held by others within the community.

Building on the recent empirical literature that attempts to establish correlation between financial returns to education and children's schooling, by using estimated returns to education as an explanatory factor in the demand for education, we find that returns to education are not important in either period. Two explanations are proposed for this finding: returns to education appear to change considerably between the years, suggesting they may vary too much for a household to include in their decision making;

and households may still be unable to respond to higher returns to education due to credit constraints.

As for the development of these parameters, we find that while the grade-for-age ratio of children have risen considerably over the years, there are no significant differences in the importance of the level of consumption, the choice of livelihood of the household, the level of education in the community or the parents' own education for children's schooling. This indicates that the role of the socio-economic standing of the household has not changed between the surveys, and hence remains an important source of educational inequality. In other words, while the development after the push towards UPE indicates that the government has been successful in raising the level of education across the board, it seems to have been less so in terms of reducing educational inequalities.

2 Background – conceptual framework

2.1 Setting

The Tanzanian educational system consists of seven years of primary schooling, followed by four years of lower secondary and two years of upper secondary. Although primary schooling is and has been formally mandatory, this has often not been complied with.

Following low enrolment rates in the 1990s, the government of Tanzania adopted the Education Sector Development Programme (ESDP) at the turn of the millennium. The first stage of this programme was the adoption of the Primary Education Development Program (PEDP), which was introduced in 2002 with the initial goal of achieving UPE by 2005 (URT 2006). The quantitative goals of the programme have largely been met, with net enrolment in primary school being up from 53 percent in 2000 to over 99 percent in 2008 (WDI online, 2010). However, quality indicators show conflicting trends. There has been increasing average pupil-to-teacher ratios and increasing drop out and repeat rates. Conversely, the textbook-to-pupil ratio has increased substantially, and resources devoted to training and material has increased. The negative indictors provide some concern that the government is repeating earlier mistakes. In the late 1970s, a similar push towards UPE temporarily increased enrolment rates, yet due to decreasing quality and low returns to education, the effects were unsustainable (Wedgwood 2007, World Bank 2010). However, recent results

indicate a better outcome this time. Students' reading and mathematics tests, arranged by UNESCO sponsored organization SACMEQ, have shown increases in students' achievements between 2000 and 2007 (SACMEQ 2005, 2010), suggesting that the Tanzanian government may have managed to sustain or even increased educational quality despite the massive increase in enrolment.

Still, for this latest policy initiative to be successful, it is likely that the economic underpinnings for demanding education will need to have changed. By improving on the returns to education, the initial success of the program may be sustained. Likewise, helping families cope with the costs of education, including both direct and opportunity costs, and by changing the norms and attitudes towards education in society, one may reach an enduring improvement in educational outcomes. We now turn to developing a conceptual framework to help us address some of these issues.

2.2 Conceptual Framework

In organising the analysis we draw on the previous theoretical and empirical literature on the demand for education. We consider three core concepts in determining the demand for education: the direct and opportunity costs, the benefits/returns, and household preferences. These concepts help in understanding the basis for the analysis and rationale for inclusion of variables, even though it may not always be easily distinguishable which of these is at work. For instance, while a parent's education is usually a very robust predictors of children's education, this may be due to the household facing lower costs of, higher returns to, or stronger preferences for education.

From a theoretical viewpoint, children's education may be seen both as an investment and as a consumption good. To the extent that utility is derived directly from education, schooling can be viewed as ordinary consumption. However, education also yields longer term returns through higher future income and non-financial benefits such as better health. When deciding on the education of children, a household will arguably consider both the consumption value derived from schooling and the longer term returns to education, which will be affected by the economic context of the household, the quality of education, and the child's innate ability. From the household's perspective, the benefits of education will be weighed against the costs that come with sending children to school, i.e. direct costs (school fees and costs of transportation, school books and uniforms, for which there may be scope for economies of scale within the household) and opportunity costs (all foregone income or production the child could

have contributed to the household had he or she not been in school).

2.2.1 Costs

The UPE has resulted in a reduction in the cost side of the equation. However, although the government has abolished fees for primary schooling, these fees were already a small part of the overall costs of education. A number of papers on Tanzanian data from the 1990s found that school attendance had more to do with opportunity costs than direct costs (Mason and Khandker, 1996; Al-Samarrai and Peasgood 1998; Al-Samarrai and Reilly, 2000; and Beegle and Burke, 2004). Therefore, despite the drop in direct costs, the opportunity cost will arguably remain an important component of costs faced by household decision makers. While our data does not allow estimations of direct costs of education, we are able to test whether the role of opportunity costs have changed during the period.

2.2.2 Returns

Compared to the literature on costs, the importance of the benefits or returns to education in educational decision making is less well documented. In recent years there has been an increased interest in establishing the effects of the returns to education on the demand for education by explicitly estimating returns by means of a Mincer wage equation, and using it as an explanatory variable. However, we are unaware of any study that has tried to use these estimates in a nationally representative setting in Africa. Gormly and Swinnerton (2004) consider an urban setting in South Africa, while most other authors have focused on India.

Notable evidence in the previous literature include Yamauchi (2007), who argues that the adoption of high yield variety crops caused a shift in returns to education in India and uses this to identify a causal effect of local returns to education on the demand for schooling. He shows that households learn about these returns from observing their neighbours – a finding also noted by Anderson et al. (2003) and Kochar (2004) in Malaysia and India respectively.

However, some authors have also noted that among credit-constrained households, this effect is often missing as households may be unable to respond to higher returns to education. Gormly and Swinnerton (2004) identify a theoretical ambiguity regarding the sign of the effect of higher returns to education on schooling demand in credit-constrained households. They show that while higher returns to education should imply a higher demand for schooling due to a substitution effect, there is also a negative

income effect stemming from the fact that a higher lifetime income may make households want to consume more today. If households are credit constrained, the way to increase consumption today may be to not send children to school. However, in their particular study of urban households in South Africa, they find support for a positive effect of returns on demand for education even among the poorest households. Contrasting evidence is found by Chambagwala (2008), who finds no effect of returns on educational demand among the poorest households in India whereas Kingdon and Theopold (2008) find evidence of a negative effect for boys among credit-constrained households in India.

Other authors have looked at the effects of school quality variables that are likely to shift the returns to education on the demand for schooling in Tanzania, but have found no or only weak links. Beegle and Burke (2004) find no support for effects of school quality on demand, while Bommier and Lambert (2000) find that the quality of Swahili teaching has some effect on the length of children's education, whereas the quality of mathematics teaching and the availability of school supplies do not. None of these variables are correlated with children's school starting age at the standard five percent significance level.

From the evidence available, there does not appear to have been any major changes in the economic context that ought to be responsible for any large shift in returns to education, and correlations between educational quality and schooling decisions have been found to be weak. Combined with widespread poverty and a likelihood of a high ratio of credit-constrained households in Tanzania, it is difficult to have a strong prior even on the sign of the effect of returns to education in Tanzania.

2.2.3 Preferences

Apart from deriving utility from future returns, households may also have a taste for education, i.e. they may derive some utility from children's education *per se*. Such preferences are likely to differ among households as they depend on private notions of educational ideals, but they may also have a common component based on community norms regarding the desirability of education. We will explore whether changes in such household preferences have altered the demand.

While households' preferences for education are generally unobserved, studies have used proxies to test their impact on educational demand. Al-Samarrai and Peasgood (1998) find that girls in polygamous households in Tanzania have a lower probability of

going to school, while there is no such effect for boys, reflecting perhaps some cultural norms that influence the parents in their decision to educate their children. Another common factor used to proxy for preferences is parents' own levels of education. However, while the educational levels of a mother and a father are likely to contain information about their preferences, they are also likely to be correlated with information they have on the benefits of schooling, and as Akabayashi and Psacharopoulos (1999) point out, more educated parents will be in a better position to help with homework, thus parents' education also acts as a complement to schooling. Using the Tanzanian data we will explore potential channels of both household preferences and community norms in determining the demand for education.

3 Empirical Strategy and Data

3.1 Data

This study uses data from two Household Budget Surveys conducted by the National Bureau of Statistics in Tanzania. Both surveys cover the whole of mainland Tanzania (i.e. they exclude Zanzibar). The first survey conducted in 2001 covered approximately 20,000 households and the second survey in 2007 covered approximately 10,000 households. Both surveys used almost identical questionnaires and followed the same methodology, yet they do not form a panel. Information was collected on household characteristics, including assets, housing and a one-month consumption diary; and on individual characteristics of all household members.

3.1.1 Dependent variable

Many children in developing countries start school at different ages and drop in and out of school, which makes it difficult to find a measure of schooling that corresponds to the actual investment made in education. Previous research has used a variety of measures of demand for education, the most common include enrolment, school attendance, number of hours spent studying, grade-for-age measures and test scores. Given the data at hand, we will use children's grade-for-age ratio as our measure of educational demand. It is constructed by dividing each child's highest grade attended by the grade the child is supposed to be in. The main advantage of this variable is that it captures information on the accumulated educational investments for a child.[3] This variable has

[3] A problem with using the commonly utilised measure 'enrolment' in Tanzania is that many children do not start school at the official starting age, and drop in and out of school frequently. This means that a

properties that makes it possible to interpret it as the share of the 'officially correct' amount of education that a child has actually achieved.[4] Within all age cohorts, it equals 0 for children who have never enrolled in school and 1 for children who started school at the right age and moved on to the next grade after each year. For all children in between, it measures the share of the officially 'correct' years of education that they have attained. Hence, for a seven-year-old the official level is equal to 1 year of education, for an eight-year-old it is 2 and so on. In other words, to interpret the size of the correlations, an increase of ten percentage points in the grade-for-age ratio is equivalent to one-tenth of a year extra education for a seven-year old, two-tenths of a year for an eight-year-old and so on.

3.1.2 Explanatory variables

For ease of description, variables are discussed at their level of measurement (individual, household, village and 'returns cluster') in relation to our three main concepts of costs, benefits and preferences.

Individual level. A number of child characteristics have been found to be important for educational attainment. Apart from gender and age, birth order effects may be influential. It is usually found that first-born children receive less education (at least in younger years), as there is often a greater need for them to stay at home and help with household chores, e.g. taking care of younger siblings. Theoretically, we therefore expect the opportunity cost for schooling to be higher for first-born children than for their siblings. Compounding this, younger siblings may receive help from older siblings with homework, thereby increasing their returns further. To account for these differences we include dummies for birth order[5].

There is also reason to believe that biological children may receive more education than non-biological children, due to different preferences or different expectations

---

child who is not presently enrolled in school may very well have had more education than one who is. Joshi and Schultz (2006) use an alternative variable related to the grade-for-age ratio by constructing z-scores of the highest grade attended within each age (and gender) group. This has the advantage of taking into account the dispersion in the grade-for-age ratio in higher cohorts. However, limitations are that its size is less straightforward to interpret and that it is not evident that the distribution of the z-scores fits the data in estimable models any better than the grade for age measure. Given the strengths and limitations of the different variables, we will use the grade-for-age ratio as our benchmark measure of education and use other measures in robustness checks.

[4] The exception would be children who have a grade-for-age ratio higher than 1. Few children have that though, and the ratio can then be (equivalently) interpreted as a multiplicative factor.

[5] It should be noted that we do not have any information on children who have moved out of the household. Hence, we cannot confirm that the oldest child in the household is also the first born. Our measure is ranked by age and should be seen as a proxy for birth order.

regarding future remittances among the household decision makers. Hence, we include dummies for each child's relationship to the household head, including being the child of the spouse or the grandchild of the household head.

Household level. There are also a number of household level factors that are likely to affect children's education. The most obvious control necessary is a measure of income. We will use the log of consumption per adult equivalent. The reason for its inclusion is that apart from being indicative of possible credit constraints hindering children from going to school, a higher consumption level should lower the marginal utility of the financial net effect of education, thereby possibly giving a higher relative weight to utility derived directly from education. We also include parents' education as these may affect both preferences and returns, as discussed before.

Variables on households' productive assets that may shift the marginal productivity of child labour and hence the opportunity costs of education are also included. These variables include the log of the value of working capital and the log of the area of land owned or used for agriculture, both measured per adult in the household.[6] While these may be good proxies for opportunity costs in households engaged in agricultural or own business, they are less so for wage earning households. Hence, we include dummies for different livelihoods of households defined from statements on the main source of cash income for the household; namely, being involved in agriculture, having an own business, or being wage earners. If households believe that their children will earn their livelihood from the same activity, these dummies will capture both differences in opportunity costs and possible differences in expected returns to education. In addition, to proxy for the other costs of sending children to school we include two distance variables: the distance to the nearest primary school in kilometres – to account for transport costs (in money or time) – and the time it takes to fetch fresh water (in hours).

Finally, we include variables on the demographics of the household: the number of children, which may affect the costs of schooling since siblings may be able to share or inherit school material; the number of adults; and a dummy for having at least one grandparent present, since grandparents may be substitutes for children in certain household chores and hence reduce the opportunity costs of schooling.

Cluster level. In the next section we define a 'returns cluster' level where we divide

[6] In order not to lose households without capital or land, we add 1 to capital and the minimum non-zero area per worker (very small) in the sample before taking logarithms, and add two dummies for having no land and no capital respectively.

the sampled villages according to their location and an urban-rural distinction. At this level we introduce our return to education variable, described in detail in the next section. Importantly, we also construct a variable for the average level of education among the adults in the 'returns cluster'. We do this as there is reason to believe that people's decisions on education may be affected by the norms of the community in which they live, as these may affect the households' preferences for education.

By measuring the average level of education in different communities, we want to capture variation in social norms regarding schooling. However, a relationship between the average level of education among adults in the community and the decisions of the households regarding their children's schooling may come about for several reasons[7]: parents may send their children to school because other parents tend to do so (what Manskie (1993) refers to as an endogenous effect); because other parents have a high level of education (an exogenous effect) or because the average level of education is correlated with other community or household characteristics that affect the educational decisions (correlated effects). Whereas the first two effects can be interpreted as representing related social norms (on sending your children to school and on the value of education respectively), the latter is potentially more problematic for our purposes. There are at least two concerns here. One issue is that a household in an area with a high average level of education may be expected to have a relatively high level of education and a relatively high level of income themselves. To deal with this, we include control variables in our estimations in order to capture such characteristics. The other concern is that areas with a high level of education may share other characteristics that relate to children's schooling, such as a relatively high educational quality. We are not able to control directly for the quality of education, but we do control for local returns to education which may capture quality differences in schools. As mentioned it can also be noted that previous research from Tanzania has found very weak, if any, evidence that the educational quality plays any major role in Tanzanian households' decisions on children's education, though this is of course not proof that quality is unimportant.

3.1.3 The return to education

To capture systematic variation in returns to education using cross sectional data, we want to group people together in a way that makes it plausible that they face similar returns within each group but different returns across groups. We do this by estimating

[7] A problem similar to what Manskie (1993) famously referred to as the 'reflection problem'.

returns to education for specific groups defined by their location, including a measure of closeness to markets, and gender. The rationale for these dividing lines is that your location, closeness to market and gender will offer different possibilities in terms of livelihoods and hence possibly differing returns. Given the geographical size of Tanzania and the many ethnic groups in the country, there will likely be significant variation in returns by region. This is supported by evidence that sectors of employment are markedly different between Dar es Salaam and other urban areas, and even more so between urban and rural areas, where agriculture is overwhelmingly predominant (National Bureau of Statistics, 2009). Poverty rates differ dramatically by region, and while they have almost halved in Dar es Salaam since 1991/92 they have changed only slightly in rural areas (National Bureau of Statistics, 2009). Taking this argument further, the economic context of different localities will differ depending on how connected an area is to wider markets. Localities close to markets will face different exposure to outside technology, different degrees of industrialisation, and different livelihood opportunities due to a potentially more diversified demand for goods and services. Likewise, different regions in the country may have different cultural contexts, be more or less connected to the world market, and have different production traditions in terms of both technology and the goods produced. Finally, the division along the gender dimension is motivated by the fact that men and women often have different traditional roles in production and hence may have very different returns to education.

Mainland Tanzania is made up of 21 regions. We divide households within each region into urban and rural, which should capture a household's closeness to markets. This gives us 42 potential groups based on region and ruralness. We call these our 'returns clusters', and in each of these clusters we will estimate returns to education for men and women separately, giving rise to 84 different rates of return.

3.2 Estimation Strategy

In order to include the observed return to education as a predictor of the demand for education, we need to carry out our analysis in two steps. In the first step we estimate the return to education and in the second step we estimate the demand for education.

3.2.1 Estimating the Return to Education

We estimate the return to education within each returns cluster by means of an estimation similar to a standard Mincerian wage regression. However, unlike previous studies of returns set in Tanzania, which use wages from wage work as the dependent

variable (see e.g. Schultz, 2004; Söderbom et al., 2006; and Al-Samarrai and Reilly, 2008), we will use consumption. We do this for several reasons. First, for most Tanzanians, wage-based estimates may be highly misleading as wage work is the exception rather than the rule, especially in rural areas. Investigating schooling for all children, we are interested in returns to education for the whole population and not only wage earners. Second, consumption fluctuates less than income as households' smooth consumption in the presence of income shocks, hence being a better measure of the household's permanent income. Third, using income data is problematic in developing countries due to the noise in its measurement, whereas consumption has the advantage of being more precisely measured.

We measure a linear effect of years of education and use this as a benchmark estimate. Acknowledging that this may not be a completely accurate description of the returns, we use two alternative measures as robustness checks: a quadratic form and two dummies for educational attainment (one for having completed primary and one for having completed secondary education).[8] Apart from the fact that a linear return gives more stable estimates, in the presence of convex or concave returns it will give us an average return based on the levels of education present in the community which seems to be a measure that should lie close to households' expected returns.

More formally, a standard Mincer style regression allowing for gender specific returns would be estimated at the individual level as:

$$\ln(consumption_i) = \alpha + X_i\beta + \gamma D_i + \delta_1 edu_i + \delta_2 edu_i D_i + \varepsilon_i \quad (1)$$

where X_i is a vector of control variables, D_i is a gender dummy, and edu_i is years of education. In our data consumption is measured at the household level though, requiring a more aggregated estimation. Allowing for differing intercepts and returns depending on cluster, and following the methodology of Kingdon and Söderbom (2007), taking means over the working adult members of the household (anyone over 15 who is not in school) would give us the equivalent model:

$$\overline{\ln(consumption_\iota)} = \alpha^c + \bar{X}_\iota\beta + \gamma^c\bar{D}_\iota + \delta_1^c\overline{edu_\iota} + \delta_2^c\overline{edu_\iota D_\iota} + \bar{\varepsilon}_\iota, \quad (2)$$

[8] There is some evidence of convex returns to education in Tanzania (Söderbom et al. 2006), suggesting we use a quadratic function. However, this implies the need to evaluate returns at a specific level of education. Given that different groups in Tanzania have very different educational attainments (e.g. the share of people with university education in rural areas is extremely low), such a measure turns out to be imprecise, as the quadratic term makes predictions shaky when evaluated far away from the actual observations. The option to identify returns to having completed different levels of education leads to multicollinearity and in some cases rather shaky estimates. Given these caveats, we use the more robust linear form.

where a bar denotes averaging over household members. In this equation, the household mean of $\ln(consumption_i)$ is still unknown, however. Replacing the average of the log of consumption with the log of average consumption introduces a small error to the dependent variable yet makes the equation easily estimable as:

(3) $$\ln(\overline{consumption_\iota}) = \alpha^c + \bar{X}_\iota\beta + \gamma^c\bar{D}_\iota + \delta_1^c\overline{edu_\iota} + \delta_2^c\overline{edu_\iota D_\iota} + \bar{\varepsilon}_\iota.$$

It is important to note here that the control variables in X_i should not include variables caused by education. For example, since a person's education will affect his or her probability of different labour market opportunities and livelihoods, we do not want to condition consumption upon that – being a farmer or being able to get a wage job, and the effects that has on income and consumption, is part of the returns to education.

Using estimates of returns from equation 3 assumes that households base their expected returns to children's education on the actual outcomes of older generations who have finished school and are working, and on that they form these expectations using the outcomes within their 'returns clusters'. This would be in line with previous research, which finds that the current state of returns within your local community indeed affects expected returns (Jensen, 2008; Yamauchi, 2007).

The most commonly noted problem of estimating returns to education is the potential existence of an ability bias in the returns equation, i.e. that people with a higher unobserved ability will also get more education, making education correlated with the error term leading to biased estimates of the returns parameters. One solution to biased estimates is to instrument for education. We argue that there are at least two reasons why instrumentation may not be a good thing for the purpose of this study. First, as Yamauchi (2007) argues, people cannot learn about returns to education by observing themselves, as there is no counterfactual outcome. Instead, they learn from others. Yamauchi shows that farmer households in India learnt about new returns to education brought about by the 'Green Revolution' by observing the actual outcomes of their neighbours. This implies that a Mincer style equation approach may better resemble the perceived returns to education than does an approach uncovering the 'true' returns. Second, the focus of the paper is not to establish the returns to education, but rather to examine how these returns affect schooling decisions. It need not matter if the returns to education are biased as long as this bias is not different between the different returns clusters.

3.2.2 Estimating the Determinants of Education

To investigate the determinants of the demand for education, using OLS we regress the grade-for-age on our explanatory variables with standard errors clustered at the returns cluster level. We include in the analysis only children aged 7-15 – children below 7 have rarely started school, and those over 15 have often moved out of the household (which could imply a serious selection bias).

The benchmark estimation of child *i*'s educational attainment will be of the form:

$$(4) \qquad grade\_for\_age_{i,t} = \alpha_t + C_{i,t}\beta_t^C + H_{i,t}\beta_t^H + V_{i,t}\beta_t^V + R_{i,t}\beta_t^R + \varepsilon_{i,t},$$

where *C*, *H*, *V*, and *R* are vectors of child, household, village and returns cluster level variables respectively, ε is a random error term, and the *t* subscript denotes survey year and is added to underscore that estimations are undertaken with year-specific parameters.

As it is not possible to send your child to school for a negative number of years, our dependent variable cannot take on values below zero. Since a linear specification can predict values below zero, and conceptually the marginal effects can be low for individuals close to 0, one suggestion would be to use Tobit rather than OLS to estimate the demand for education. However, a Tobit estimation requires that the dependent variable can take on values close to the limit (i.e. close to zero) or the results will be biased. In our sample, the majority of children have grade-for age ratios between 0 and 1. Among the youngest children, where censoring is most common, values close to zero are not possible: for seven-year-olds the lowest non-zero value attainable is 1, for 8-year-olds it is 0.5, etc. There is also a growing literature that concern over functional form is less important than correct identification. Angrist and Pischke (2009) show that the interpretation of marginal effects present no special challenges whether the dependent variable is binary, non-negative or continuously distributed. Instead they argue that once output from nonlinear models are converted into marginal effects the differences in the OLS and nonlinear models are indistinguishable, concluding that the complexities that arise from nonlinear models outweighs the advantages of using standardized OLS estimates. We therefore present our OLS estimates in the main text, and the highly similar results from equivalent Tobit estimations in the Appendix.

4 Results

Table 1 presents summary statistics of our variables in the samples of children for both

2001 and 2007. Before turning to the demand for education estimations though, we will start by looking at the returns to education.

4.1 Estimation of returns to education

Table 2 shows the results from estimating equation 3. We regress the log of consumption per adult equivalent on age and age squared to capture life-cycle and experience effects, and the dependency ratio to control for households smoothing incomes over time, hence reporting higher consumption at times when there are many children in the household. There is an initially positive but decreasing effect of age, and a positive parameter on the dependency ratio. This is in line with expectations, and the results from both survey years are reassuringly similar. We estimate returns specific to each cluster and, as is standard in the literature, allow them to differ by gender, giving us 84 different returns. Since it is not feasible or useful to present all 84 returns, Table 2 only reports the estimated returns for region 1 (Dodoma region).

Table 3 summarises the estimated returns to education by gender and location. The returns to education are, on average, similar between urban and rural areas and between men and women, yet tend to be slightly higher for men and in urban areas. F-tests of all 84 estimated returns being equal is firmly rejected at the one percent level in both the 2001 and 2007 samples, and hence we conclude that there is strong statistical evidence that the returns clusters have differing returns to education. However, the correlation coefficient between the clusters' returns in 2001 and 2007 is only 0.13, with a p-value of 0.25, indicating a low level of correlation over the years. Hence, it seems that the pattern of returns to education may have changed over time, also suggesting that the present return to education in a community may be a poor predictor of future returns. If households realise this, it seems to make little sense for them to make use of the present return to education in their schooling decisions. Whether they do or not is a question for the empirical analysis, to which we now turn.

4.2 Determinants of the demand for education

Figure 1 shows the share of children currently enrolled in school by age cohort in 2001 and 2007, and Figure 2 shows the average grade-for-age ratios at different ages for each year. Both enrolments and grade-for-age ratios increased from 2001 to 2007, and especially so for the youngest children. This is expected as the youngest children are most likely to have been fully affected by the measures taken by the government in the PEDP, and were the ones lagging furthest behind prior to the programme.

Following our conceptual framework, Table 4 presents the results of estimations on key returns, costs and preferences variables separately, both with and without control variables, for 2001 and 2007. Column 5 reports the differences in the coefficients between the two years (for the estimations with controls), and the statistical significance of these differences. For ease of presentation, the coefficients on controls are only reported in the full estimation in Table 5.

4.2.1 Returns to education

Beginning with returns in Table 4, with and without controls, we report that returns to education do not seem to have a statistically significant impact on schooling in either year. Nor is the difference between these coefficients significant. The coefficients are not only statistically insignificant, given the small standard deviation of returns, they are also very small.

We argue that this finding is perhaps not surprising in this setting. First, as noted in Section 2, the effect of higher returns to education is theoretically ambiguous in the presence of credit constraints (Gormly and Swinnerton, 2004). As many Tanzanian households are poor and can be believed to be credit constrained, the absence of an effect is consistent with both theory and previous research, which has tended to find insignificant or negative returns among the poorest households (Chambagwala, 2008; Kingdon and Theopold, 2008). Second, the variation in returns in the 2000s may make it difficult for households to use this information when making decisions regarding schooling. Hence, it is possible that people's expectations regarding returns are formed with respect to other information, such as children's innate ability or the likelihood of migrating to areas with different returns.

4.2.2 Costs of education

Our most direct measure of costs of education, the distance to the nearest primary school, is negatively and statistically significantly correlated with schooling at the one percent level in all estimations of Table 4. This result remains in Table 5, which presents estimations for the full set of variables (but drops the returns to education which turned out to be insignificant and which is itself estimated, biasing its standard error downwards[9]). However, the coefficient is fairly small – the 2007 estimate of about -0.002 in Table 5 indicates that an increase in the distance to school by one standard

[9] There are alternative ways of accounting for the fact that the return variable is estimated, but as we shall see the return variable adds little of value to the estimations, so it is more efficient to drop it.

deviation of 7.57 kilometres decreases the education attained by just 1.4 percentage points of the officially correct level. The impact is therefore limited, which is in line with other research; Filmer (2004), investigating 21 poor countries (including Tanzania), found that while the coefficient of distance to school is generally negative and significant, the size is typically small in relation to other determinants. On the other hand, the impact was significantly larger in 2001, when the point estimate was about 150 percent greater, implying that living 7.57 kilometres further away from the closest primary school was predicted to decrease the education attained by 3.6 percentage points. Hence, it seems that the PEDP may have been successful in promoting children's education among those living furthest away from schools.

Opportunity costs of education are captured by a number of variables: those that involve household chores, those that free up time for children, and household demographics. The most common children's household chore is collecting water. The time it takes to fetch fresh water enters as negative and statistically significant at the one percent level in 2001 both with and without controls in Table 4. For 2007, the parameters are similar but statistically insignificant. Table 5 yields similar results, with an insignificant parameter for 2007 but also an insignificant difference in parameters. Like distance to school, the coefficients for both years are small and has little impact; the (larger) 2001 parameter indicates that an increase in the water fetching time by one hour was associated with a decrease in educational attainment by 1.8 percentage points of the officially correct level.

Productive assets are also assumed to increase the opportunity costs of schooling by raising the marginal productivity of child labour. In line with this, the log of land owned per working adult enters negatively and significantly for 2001 without controls in Table 4, but loses its statistical significance with controls in 2007. The log of working capital per working adult enters positively and significantly in both years. However, in the full regressions of Table 5, the parameters of both land and capital ownership are statistically insignificant and insignificantly different between the survey years. It may be noted here that while such measures often come out as significant in the literature, the theoretical link to schooling is ambiguous. In itself, working capital may enhance the marginal productivity of child labour, but may also replace child labour in production. For instance a household that owns an ox plough or pesticide sprayer may have less of a need for child labour in agriculture. Additionally, with more land the schooling of children may be more important for future production.

Related to the measures of productive assets are the dummies indicating the household's main source of cash income. In 2007, children in households relying on wage earning activities are estimated to have about 3.3 percentage points higher grade-for-age ratio than those relying on agriculture, while the parameter for operating an own business is much smaller at 1.7. Both parameters were statistically insignificant in 2007. That of being a wage earner was both larger and statistically significant in 2001, when being the child of a wage-earner indicated an increased schooling of approximately 8 percentage points of the correct level. This parameter is in line with the hypothesis that wage earners have less of a need for child labour in income-generating production. Children in other households may often have to help out in the fields or in the household business. The exact causal mechanism is unclear however. There is evidence that wage work in Tanzania has higher returns to education than own employment (Al-Samarrai and Reilly, 2008), and it is widely believed that the returns in agriculture are even lower. Hence, it seems probable that wage earners may also have different expectations regarding the returns to education. Column 3 indicates that even though the point estimates were larger in 2001, possibly implying a decreased importance of family livelihood as a determinant for children's schooling, the differences between the surveys are again statistically insignificant.

In Table 5, we also see that birth order is expected to impact demand, which we interpret as an effect of first-born children often having to look after younger siblings. This parameter is negative and statistically significant in 2007 but not in 2001; the disadvantage of being the eldest child in the family has grown significantly over the years, indicating a level of education almost 4 percentage points (of the correct level) lower for first-borns than for younger children in 2007.

Finally, as expected the parameter of household consumption is positive and statistically significant in both years. The sizes of the parameters are fairly small but similar in 2001 and 2007, implying a modest but sustained impact on the demand for primary education; the standard deviation of our consumption measure is approximately 0.54 in both years, which implies a change in the grade-for-age ratio of approximately 3 percentage points. This can be seen as support for the hypothesis that credit constraints are at play. It also suggests that preferences for education are less important relative to the marginal utility derived from child labour at lower income levels.

4.2.3 Preferences

Returning to Table 4, we next consider preferences for education. The average level of education of the other households within the cluster is, as discussed, included as a proxy for shared common preferences for education in the community (such as social norms regarding education). The variable is large, positive and significant across all estimations. Introducing controls and other variables in the full estimation its size is about halved, yet it remains statistically significant at the one percent level, and predicts a 4.7 percentage point higher educational attainment in 2007 for each extra year of average education within the cluster. As the standard deviation of the average years of education is about 1.2 and the minimum and maximum levels are 3.2 and 7.5 years of education respectively, this implies a sizeable effect if interpreted as a causal relationship. This lends support to the hypothesis that people are affected by the standard norm regarding sending children to school. Turning to Table 5, the parameter keeps both its economic and statistical significance in the full estimations, and is still very similar in 2007 and 2001. This suggests that the government's push for UPE does not seem to have affected areas of lower levels of education more than it has areas with higher levels, other than possibly through other variables included. Hence, it seems that while children in all communities are now expected to go to primary school to a larger degree, the relative disadvantage of living in low-education settings is sustained. As discussed in section 2, we cannot rule out that some other factors than community norms are also at play. However, such factors have been controlled for when possible. To the extent that there are persistent differences in the quality of education between clusters, this effect may be picked up by the return to education variable, and persistent supply of education effects should be controlled for with the distance to school included. But there could also be informational effects at play, as people living in clusters with higher educational levels may have a clearer picture of the benefits of education or may simply believe they have imperfect information and hence copy others who may have better information. If so, the stability of the parameter of the average level of education would indicate that any informational campaigns do not seem to have reached out to areas with lower education to any larger extent than to those with higher education.

Moving on to the own parents' educations in Table 5, the parameters of the educations of the child's mother and father may also be interpreted as effects of differing preferences. Parents with higher education may have chosen to get an

education because they have higher preferences for education, or they have developed higher preferences due to their education. Either way, the parameters on parents' educations are large, positive and stable over the years. The impact of the mother's education tend to be higher than that of father's; in 2007 a child with a mother with primary education is predicted to have approximately 9.5 percentage points higher grade-for-age ratio than one with an uneducated mother. This is a slightly stronger effect than that of having a father with at least secondary education. Yet, the impact of having a mother with secondary education is even greater (double that of primary). These parameters are not statistically significantly different in 2001 and 2007, but somewhat larger for the father and lower for the mother in the earlier survey. The difference between the mother's and the father's education may to some degree be a selection effect; women tend to have a lower level of education, thus their educational attainment may be a stronger signal of their preferences. However, it also follows the patterns in previous research.

There are also strong negative effects of being the child of the spouse of the head of household (but not of the head) or some other related or unrelated child living in the household. For both groups of children, the expected grade-for-age ratio is much lower than for biological children to the head of the household; in 2007 a child of the spouse had a predicted schooling 17 percentage points of the correct level lower than that of a biological child. This is likely an effect of both preferences, and that the household expects to receive less remittance from non-biological children, and therefore does not see investment in their human capital as important. The latter may not be as applicable for the children of the spouse, as for completely unrelated children. In contrast to the large parameter of 2007, the 2001 parameter of being a child of the spouse did not even imply a 1 percentage point difference to biological children of the household head, indicating that non-biological children have actually fallen behind the others over time.

Finally, in line with other research the coefficient for being male is negative, large and significant. This corresponds to previous research from Tanzania (Bommier and Lambert, 2000), which has found that boys tend to start school later than girls (hence the negative parameters) yet ultimately receive more education (which would show up in later years). This may reflect parents' preferences (discriminatory preferences result in less education for girls), higher returns to pre-school experience from the household's economic activities for boys (hence higher opportunity costs) and the existence of a bride price (higher returns to early education for girls). The negative effect of being a

boy is larger in 2007 than in 2001, though the difference is statistically insignificant.

5 Conclusion

This study set out to provide new evidence on the determinants of the demand for education in Tanzania and to see how these have developed during the push for UPE in the last decade. We have done so by utilising data from two household budget surveys conducted in 2001 and 2007 that cover mainland Tanzania. The focus has been on determining the importance of variation in the costs of education, the returns to education, and household preferences for the educational attainment of children of primary school age.

Drawing on the existing theoretical and empirical literature, we include determinants of education that represent costs, benefits and preferences. There is evidence that direct and opportunity costs of children's education are important factors, and there is some evidence that living close to school is significantly less important in 2007 than in 2001. However, we also find that after the introduction of the PEDP, the role of family livelihoods have not changed significantly. Likewise, the level of consumption is still a significant determinant of demand. This indicates that despite a general increase in schooling, structural differences in educational attainment due to the economic standings of the households have remained largely intact over the period.

The paper estimates local returns to education which is then used as an explanatory factor in the demand for education estimations. In contrast to a number of recent studies we find no evidence of effects of local returns to education on the demand for schooling. We argue that the decisions on primary education seem to be dominated by other factors than variations in local returns, and offer two explanations; households may be subject to credit constraints, which we see some evidence of with the size of the coefficient on consumption and in the importance of being a wage earner or involved in business, and we also note that local returns are volatile between periods, possibly making it difficult for households to use when forming expectations about future returns.

We also find several factors thought to represent household preferences for education to be important, including determinants at both the household and community levels. The average level of education in the region of the household, indicative of some common preferences in the community such as social norms regarding education, has a

statistically significant and economically important predictive power with respect to children's education, even when controlling for parents' own education. Moreover, this disadvantage of living in an area with a lower level of education among the adults has been remarkably stable. In line with other research, we also find that some of the strongest determinants of educational demand come from parents' characteristics. The relationship of the child to the household head, together with the effect of parents' education (especially that of the mother), are strong predictors of educational demand. The evidence suggests that children born to well-educated parents who remain in the custody of their biological mother and father have significantly higher levels of education, and that this is just as important in 2007 as it was in 2001.

The Tanzanian government seem to have been successful in its push towards universal primary education, but the sustained importance of consumption, livelihoods, the level of education in the community, as well as the parents' own education, indicate that the role of the socio-economic standing of the household has not changed significantly between the surveys, and hence remains an important source of educational inequality. In this light, while a development of attitudes in favour of children's education may bring positive feedback effects on the education of coming generations as future parents will be better educated, history has shown that there is also a risk that such progress may regress in the absence of substantial changes in the economic context in which the household takes its decision. While the present study find no significant correlation between the observed local financial returns to education and children's schooling, there is of course still a risk that the accomplishments achieved during the PEDP may not be sustainable and that the experience of the 1970s therefore may be repeated.

6 References

Akabayashi, H. and Psacharopoulos, G. (1999). 'The Trade-off between Child Labour and Human Capital Formation: A Tanzanian Case Study', *Journal of Development Studies*, Vol. 35, No.5, 1999.

Al-Samarrai, S. and Peasgood, T. (1998). "Educational Attainments and Household Characteristics in Tanzania', *Economics of Education Review*, Vol. 17, No. 4, 1998.

Al-Samarrai, S. and Reilly, B (2000). 'Urban and Rural Differences in Primary School Attendance: an Empirical Study for Tanzania', *Journal of African Economies*, Vol. 9, No. 4, 2000.

Al-Samarrai, S. and Reilly, B. (2008). 'Education, Employment and Earnings of Secondary School and University Leavers in Tanzania: Evidence from a Tracer Study', *Journal of Development Studies*, Vol. 44, No. 2, 2008.

Anderson, K. H., King, E. M. and Wang, Y. (2003). 'Market Returns, Transfers and Demand for Schooling in Malaysia, 1976-89', *Journal of Development Studies*, Vol. 39, No. 3, 2003.

Angrist, J. D. and Pischke, J. (2009) *Mostly Harmless Econometrics – An Empiricist's Companion*, Princeton University Press, Princeton.

Beegle K. and Burke K. (2004). 'Why Children Aren't Attending School: The Case of Northwestern Tanzania', *Journal of African Economies*, Vol. 13, No. 2, 2004.

Bommier A. and Lambert S. (2000). 'Education Demand and Age at School Enrollment in Tanzania', *The Journal of Human Resources*, Vol. 35, No. 1, 2000.

Chambagwala R. (2008). 'Regional Returns to Education, Child Labour and Schooling in India', *Journal of Development Studies*, Vol. 44, No. 2, 2008.

Gormly S. and Swinnerton K., (2004). 'The Effect of Adult Returns to Schooling on Children's School Enrollment, Theory and Evidence from South Africa', Mimeo.

Filmer, D. (2004) 'If You Build It, Will They Come? School Availability and School Enrollment in 21 Poor Countries', World Bank Policy Research Working Paper 3340, June 2004.

Joshi, S and Schultz, T. P. (2006) 'Family Planning as an Investment in Development: Evaluation of a Program's Consequences in Matlab, Bangladesh', unpublished paper, mimeo.

Jensen, R. (2010). 'The (Perceived) Returns to Education and the Demand for Schooling', *The Quarterly Journal of Economics*, May 2010.

Kingdon, G. and Söderbom, M. (2007). 'Education, Skills, and Labor Market Outcomes: Evidence from Pakistan', mimeo.

Kingdon G. and Theopold N. (2008). 'Do returns to education matter to schooling participation? Evidence from India', *Education Economics*, Vol. 16, No. 4, 2008.

Kochar A. (2004). 'Urban influences on rural schooling in India', *Journal of Development Economics* No. 74, 2004.

Manskie, C. F. (2003). 'Identification of Endogenous Social Effects: The Reflection Problem', *The Review of Economic Studies*, Vol. 60, No. 3, July 1993.

Mason A. and Khandker S. (1996). 'Measuring the Opportunity Costs of Children's Time in a Developing Country: Implications for Education Sector Analysis and Interventions', Human Capital Development Working Papers, No. 72, 1996.

National Bureau of Statistics (2009). 'Household Budget Survey 2007', final report, mimeo

Orazem P. and King E. M. (2007). 'Schooling in Developing Countries: The Roles of Supply, Demand and Government Policy', Iowa State University Working Paper, No. 07019, 2007.

Rosenzweig, M. R. (2010). 'Microeconomic Approaches to Development: Schooling, Learning, and Growth', Yale Economic Growth Center Discussion Paper No. 985, 2010.

SACMEQ (2005). 'The SACMEQ II Project in Tanzania: A Study of the Conditions of Schooling and the Quality of Education', SACMEQ Educational Policy Research Series, Tanzania Working Report, available online at: http://www.sacmeq.org/downloads/sacmeqII/Tanzania.zip

SACMEQ (2010). 'SACMEQ III Project Results: Pupil achievement levels in reading and mathematics', Working Document, No 1, available online at: http://www.sacmeq.org/downloads/sacmeqIII/WD01_SACMEQ_III_Results_Pupil_Achievement.pdf

Schultz T. P. (2004). 'Evidence of Returns to Schooling in Africa from Household Surveys: Monitoring and Restructuring the Market for Education', *Journal of African Economies*, Vol. 13, AERC Supplement 2, 2004.

Söderbom M., Teal, F., Wambugu, A. and Kahyarara, G. (2006). 'The Dynamics of Returns to Education in Kenyan and Tanzanian Manufacturing', *Oxford Bulletin Of Economics And Statistics*, Vol. 68, No. 3, 2006.

United Republic of Tanzania (URT) (2006), 'Primary Education Development Programme II', available online at: http://moe.go.tz/PDF/PEDP%20II%20Final%20Doc.pdf

Wedgewood R. (2007). 'Education and poverty reduction in Tanzania', *International Journal of Educational Development*, Vol. 27, 2007.

WDI online 2010, The World Bank website. Accessed 07-10-2010. http://data.worldbank.org/indicator

World Bank (2010). 'Project Performance Assessment Report Tanzania', Report No. 55383, June 28, 2010. Available online at: http://www-wds.worldbank.org/external/default/WDSContentServer/WDSP/IB/2010/07/30/000334955_20100730033551/Rendered/PDF/553830PPAR0p121ficial0use0only01061.pdf

Yamauchi, F. (2007) 'Social learning, neighborhood effects, and investment in human capital: Evidence from Green-Revolution India' *Journal of Development Economics*, Vol. 83, 2007.

Figures and Tables

Table 1. Weighted summary statistics of children in the demand for education samples.

	2001				2007			
Variable	Mean	Sd	Min	Max	Mean	Sd	Min	Max
grade-for-age	0.39	0.42	0	3	0.66	0.41	0	3
Return to edu	0.06	0.02	0.02	0.18	0.05	0.03	-0.05	0.16
Ln area per worker	-0.35	2.71	-7.31	6.21	-0.41	2.90	-7.17	4.67
No area (dummy)	0.13	0.43	0	1	0.15	0.48	0	1
Ln capital per worker	10.10	4.02	0	16.82	10.94	4.75	0	20.44
No capital (dummy)	0.07	0.33	0	1	0.06	0.38	0	1
Distance to school (km)	2.05	3.64	0	130	3.45	7.57	0	300
Time to water (hrs)	0.41	0.64	0	20	0.66	0.83	0	8.1
Mean cluster education	4.85	1.09	3.51	7.64	4.77	1.36	3.04	7.51
Ln consumpt. per ad.eq.	9.07	0.54	7.43	11.00	9.78	0.54	8.27	11.61
Wage earner	0.08	0.40	0	1	0.14	0.43	0	1
Own business	0.12	0.42	0	1	0.13	0.41	0	1
Other livelihood	0.10	0.34	0	1	0.08	0.30	0	1
Male	0.50	0.50	0	1	0.51	0.50	0	1
Oldest child	0.31	0.45	0	1	0.33	0.47	0	1
Other child	0.09	0.32	0	1	0.09	0.30	0	1
Child of spouse	0.02	0.15	0	1	0.03	0.15	0	1
Grandchild of head	0.12	0.32	0	1	0.13	0.34	0	1
No. of kids	4.48	2.52	1	33	4.29	2.13	0	16
Grandparent present	0.20	0.40	0	1	0.18	0.39	0	1
No. of adults	3.12	1.81	0	16	2.85	1.54	1	21
Mother's edu primary	0.44	0.50	0	1	0.49	0.50	0	1
Mother's edu secondary	0.03	0.28	0	1	0.03	0.25	0	1
Father's edu primary	0.38	0.49	0	1	0.45	0.50	0	1
Father's edu secondary	0.06	0.36	0	1	0.06	0.32	0	1
Rural	0.82	0.49	0	1	0.77	0.49	0	1

Table 2. OLS Returns: Estimations for log consumption per adult equivalent

Coefficient (t value)	(1) 2007	(2) 2001
Dependency ratio	0.128***	0.132***
	(-0.004)	(-0.006)
Age	0.026***	0.030***
	(-0.002)	(-0.002)
Age squared	-0.000***	-0.000***
	(0.000)	(0.000)
*Male*Urban*Region1*	1.308***	-0.343
	(-0.222)	(-0.313)
*Male*Urban*Region1*Education*	0.067***	0.110***
	(-0.011)	(-0.026)
*Male*Rural*Region1*	1.128***	-0.15
	(-0.215)	(-0.288)
*Male*Rural*Region1*Education*	0.047***	0.059***
	(-0.012)	(-0.02)
*Female*Urban*Region1*	1.421***	0.024
	(-0.212)	(-0.296)
*Female*Urban*Region1*Education*	0.060***	0.058***
	(-0.009)	(-0.021)
*Female*Rural*Region1*	1.166***	-0.300
	(-0.210)	(-0.285)
*Female*Rural*Region1*Education*	0.064***	0.026*
	(-0.01)	(-0.014)
Observations	19931	9366
R-squared	0.28	0.34

Note: Return to education is estimated specific to each combination of urban/rural, region and gender. All variables are household averages. Due to space limitations, the estimation shows estimated returns to education and intercepts for men and women in urban and rural areas of region 1 (Dodoma) only; summary statistics of the returns in all 21 regions can be found in Table 2. Robust standard errors in parentheses. * significant at 10%,** significant at 5%, *** significant at 1%.

Table 3. Summary statistics of the estimated returns to education, by gender and urban/rural

	2001				2007			
Cluster	Mean	St.Dev.	Min.	Max.	Mean	St.Dev.	Min.	Max.
Female Rural	0.054	0.013	0.021	0.076	0.024	0.029	-0.050	0.061
Female Urban	0.068	0.011	0.040	0.092	0.063	0.029	0.007	0.116
Male Rural	0.067	0.031	0.045	0.183	0.050	0.021	0.022	0.097
Male Urban	0.070	0.014	0.041	0.096	0.083	0.037	0.003	0.163
All	0.065	0.020	0.021	0.183	0.055	0.036	-0.050	0.163

Figure 1. Children's average enrolment in 2001 and 2007, by age.

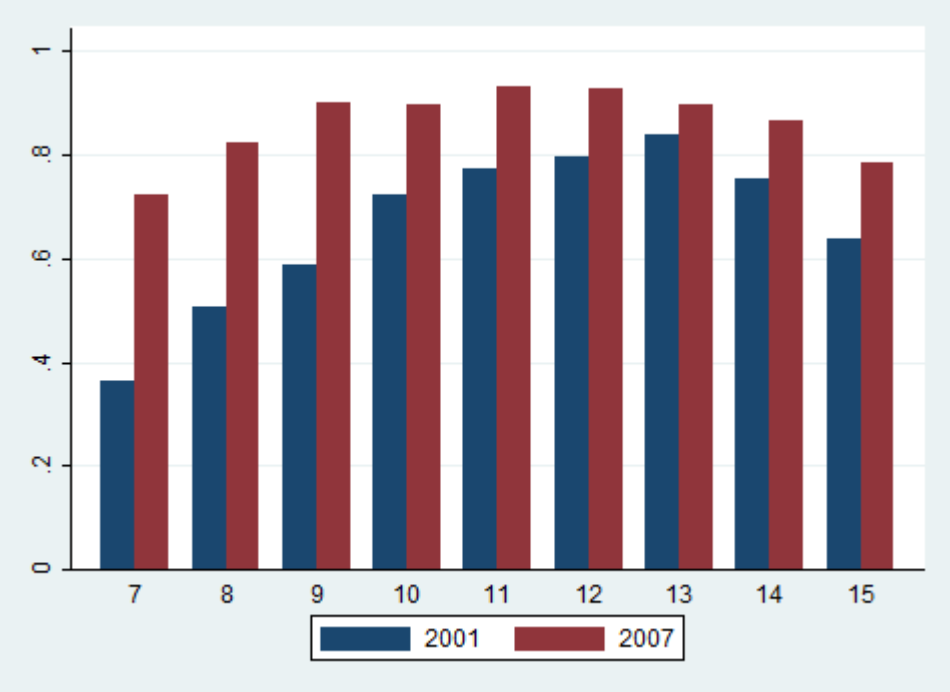

Figure 2. Children's average grade-for-age ratio in 2001 and 2007, by age.

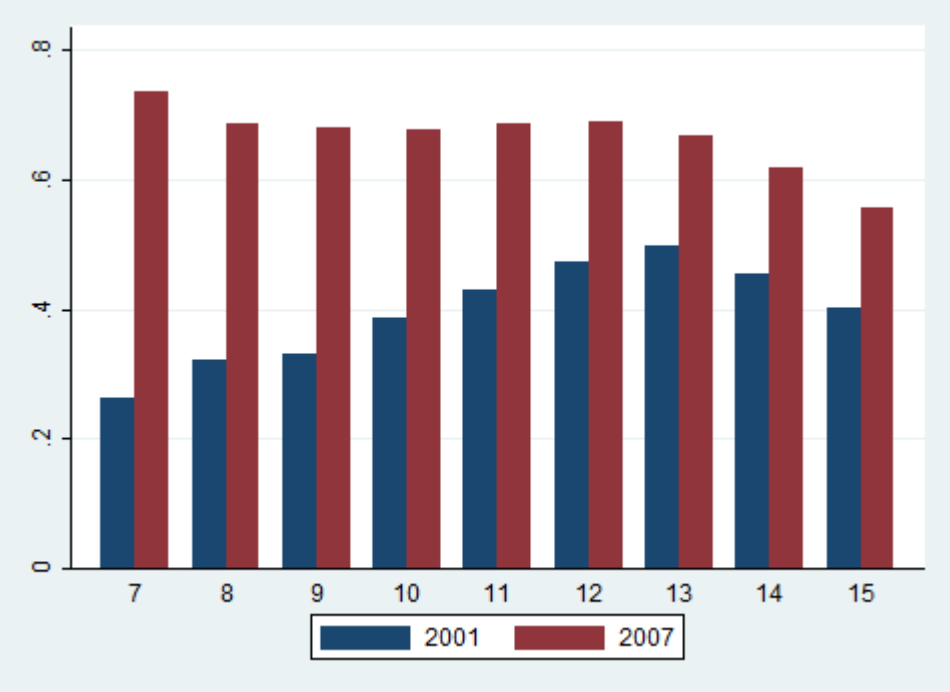

Table 4. Results of OLS estimations of children's grade-for-age ratios in 2001 and 2007.

	(1) 2001	(2) 2001	(3) 2007	(4) 2007	(5) Diff
Returns					
Return to education	0.157	-0.613	0.141	-0.175	0.438
	(1.031)	(0.485)	(0.355)	(0.312)	(0.601)
Controls	no	yes	no	yes	yes
Observations	21,298	21,298	8,549	8,549	29,847
R-squared	0.038	0.137	0.013	0.104	0.213
Costs					
Ln area per worker	-0.0202**	-0.0112	-0.000942	0.00174	0.0130
	(0.00874)	(0.00723)	(0.00657)	(0.00780)	(0.0102)
Ln capital per worker	0.0147**	0.0107	0.0136***	0.00805**	-0.00263
	(0.00723)	(0.00643)	(0.00415)	(0.00400)	(0.00663)
Distance to school (km)	-0.00866***	-0.00552***	-0.00431***	-0.00234***	0.00318**
	(0.00160)	(0.00130)	(0.00109)	(0.000662)	(0.00126)
Time to water (hrs)	-0.0299***	-0.0186***	-0.0225*	-0.00622	0.0124
	(0.00684)	(0.00610)	(0.0129)	(0.0119)	(0.0144)
Controls	no	yes	no	yes	yes
Observations	21,298	21,298	8,549	8,549	29,847
R-squared	0.068	0.139	0.046	0.108	0.124
Cluster mean education					
Average education among adults (yrs)	0.0725***	0.0479**	0.0735***	0.0469***	-0.00104
	(0.00925)	(0.0185)	(0.00814)	(0.0130)	(0.0196)
Controls	no	yes	no	yes	yes
Observations	21,298	21,298	8,549	8,549	29,847
R-squared	0.074	0.143	0.055	0.109	0.217

Note: OLS estimations of children's grade-for-age separately on returns, opportunity costs and the cluster average of years of education among adults respectively, with and without controls for 2001 and 2007. Estimation 5 estimates the difference in coefficients between the two samples. Controls included in estimations 2, 4 and 5 are those included in the full estimation in Table 5 and are left out due to space limitations. All costs estimations also include dummies for having no land and no capital respectively. Standard errors clustered at the 'returns cluster' level. * significant at 10%; ** significant at 5%, *** significant at 1%.

Table 5: Full (weighted) OLS coefficients of the demand for education 2001 and 2007 – children's grade for age

Coefficient	(1) 2001	(2) 2007	(3) Difference
Costs			
Ln area per worker	-0.00697	0.00743	0.0144
	(0.00721)	(0.00820)	(0.0109)
Ln capital per worker	0.00820	0.00591	-0.00229
	(0.00614)	(0.00402)	(0.00732)
Distance to school (km)	-0.00473***	-0.00189***	0.00284**
	(0.00123)	(0.000681)	(0.00140)
Time to water (hrs)	-0.0176***	-0.00250	0.0151
	(0.00611)	(0.0123)	(0.0136)
Mean cluster education			
Average years of education among adults in cluster	0.0393**	0.0416***	0.00226
	(0.0197)	(0.0148)	(0.0246)
Child characteristics			
Male	-0.0495***	-0.0708***	-0.0214
	(0.0179)	(0.0170)	(0.0246)
Oldest child	0.0138	-0.0388***	-0.0526**
	(0.0151)	(0.0135)	(0.0202)
Other child	-0.118***	-0.0824***	0.0355
	(0.0191)	(0.0184)	(0.0265)
Child of spouse	-0.00696	-0.172***	-0.165***
	(0.0415)	(0.0343)	(0.0537)
Grandchild of HH head	0.0435	0.0282	-0.0154
	(0.0358)	(0.0452)	(0.0574)
Household characteristics			
Ln consumption per adult equivalent	0.0664***	0.0509***	-0.0155
	(0.0138)	(0.0136)	(0.0193)
Wage earner	0.0798**	0.0326	-0.0472
	(0.0311)	(0.0197)	(0.0367)
Own business	0.0401	0.0169	-0.0232
	(0.0285)	(0.0187)	(0.0340)
Other livelihood	0.00436	-0.00144	-0.00580
	(0.0223)	(0.0213)	(0.0307)
Grandparent present	0.0205	0.0449	0.0243
	(0.0278)	(0.0412)	(0.0496)
No of adults	0.00176	-0.0214***	-0.0232***
	(0.00412)	(0.00589)	(0.00716)
No of kids	-0.0108***	-0.00517	0.00562
	(0.00288)	(0.00355)	(0.00456)
Mother's edu - primary	0.0847***	0.0950***	0.0103
	(0.0167)	(0.0123)	(0.0207)
Mother's edu - secondary	0.129***	0.182***	0.0525
	(0.0342)	(0.0312)	(0.0461)
Father's edu - primary	0.0410***	0.00832	-0.0327
	(0.0151)	(0.0141)	(0.0207)
Father's edu - secondary	0.139***	0.0811***	-0.0584
	(0.0254)	(0.0287)	(0.0382)
Rural	0.0306	0.00172	-0.0289
	(0.0414)	(0.0305)	(0.0512)
Constant	-0.489***	-0.135	0.355*
	(0.162)	(0.150)	(0.218)
Observations	21,298	8,549	29,847
R-squared	0.148	0.111	0.221

Standard errors (in parenthesis) clustered at the 'returns cluster' (by year, in Column 3) level. Additional controls included: dummies for age, having no land and having no capital. * significant at 10%; ** significant at 5%, *** significant at 10%. Additional variables: child age dummies.

APPENDIX

Table A1: Tobit estimation coefficients (Col 1-3) and their implied marginal effects at the mean of the other independent variables (Col 4-5)

	2001	2007	Diff	Mfx 2001	Mfx 2007
Ln area per worker	-0.0145	0.00970	0.0242*	-0.00980	0.00893
	(0.0107)	(0.00980)	(0.0146)	(0.00726)	(0.00900)
Ln cap. per worker	0.0127	0.00677	-0.00597	0.008627	0.00623
	(0.00929)	(0.00468)	(0.00993)	(0.00627)	(0.00432)
Distance to school	-0.00987***	-0.00282***	0.00705**	-0.00668***	-0.00259***
	(0.00317)	(0.00103)	(0.00298)	(0.00214)	(0.000951)
Time to water	-0.0329***	-0.00108	0.0319	-0.0223***	-0.000992
	(0.0117)	(0.0155)	(0.0204)	(0.00799)	(0.0142)
Mean cluster edu	0.0570**	0.0479***	-0.00907	0.0386**	0.0441***
	(0.0288)	(0.0174)	(0.0335)	(0.0196)	(0.0156)
Ln cons. p. ad. eq.	0.105***	0.0625***	-0.0427	0.0712***	0.0575***
	(0.0200)	(0.0158)	(0.0269)	(0.0135)	(0.0145)
Wage earner	0.0982**	0.0356	-0.0626	0.0692**	0.0329
	(0.0444)	(0.0240)	(0.0529)	(0.0325)	(0.0223)
Own business	0.0508	0.0183	-0.0325	0.0350	0.0169
	(0.0411)	(0.0216)	(0.0426)	(0.0289)	(0.0200)
Other livelihood	0.00241	0.00353	0.00112	0.00163	0.00325
	(0.0351)	(0.0248)	(0.0468)	(0.0238)	(0.0228)
Male	-0.0658**	-0.0830***	-0.0172	-0.0445**	-0.0764***
	(0.0280)	(0.0200)	(0.0319)	(0.0191)	(0.0183)
Oldest child	0.0143	-0.0479***	-0.0621**	0.00968	-0.0439***
	(0.0224)	(0.0161)	(0.0268)	(0.0152)	(0.0147)
Other child	-0.183***	-0.104***	0.0793*	-0.114***	-0.0937***
	(0.0324)	(0.0240)	(0.0450)	(0.0189)	(0.0211)
Child of spouse	-0.0204	-0.216***	-0.196**	-0.0137	-0.190***
	(0.0730)	(0.0475)	(0.0813)	(0.0485)	(0.0390)
Grandchild	0.0763	0.0459	-0.0303	0.0532	0.0425
	(0.0571)	(0.0565)	(0.0775)	(0.0411)	(0.0525)
No. of kids	-0.0152***	-0.00484	0.0104	-0.0103***	-0.00445
	(0.00442)	(0.00429)	(0.00653)	(0.00305)	(0.00395)
Grandp. present	0.0293	0.0455	0.0163	0.0200	0.0421
	(0.0439)	(0.0511)	(0.0723)	(0.0301)	(0.04759)
No. of adults	-0.00305	-0.0293***	-0.0262***	-0.00207	-0.0270***
	(0.00614)	(0.00783)	(0.00978)	(0.00416)	(0.00716)
Mother's edu - primary	0.141***	0.111***	-0.0295	0.0959***	0.102***
	(0.0263)	(0.0148)	(0.0291)	(0.0185)	(0.0136)
Mother's edu - secondary	0.186***	0.199***	0.0123	0.137***	0.188***
	(0.0461)	(0.0350)	(0.0580)	(0.0358)	(0.0335)
Father's edu - primary	0.0725***	0.0136	-0.0590**	0.0495***	0.0125
	(0.0233)	(0.0169)	(0.0258)	(0.0162)	(0.0155)
Father's edu - secondary	0.201***	0.0941***	-0.107**	0.148***	0.0877***
	(0.0333)	(0.0315)	(0.0496)	(0.0258)	(0.0297)
Rural	0.0540	0.00437	-0.0497	0.0359	0.00402
	(0.0593)	(0.0344)	(0.0659)	(0.0389)	(0.0316)
Constant	-1.106***	-0.314*	0.792***		
	(0.227)	(0.173)	(0.253)		
Sigma	0.522***	0.450***	0.522***		
	(0.0127)	(0.00957)	(0.0127)		
N	21,298	8,549	29,847	21,298	8,549

Note: Standard errors (clustered at returns cluster level) in parentheses. *** p<0.01, ** p<0.05, * p<0.1