Household Composition and Schooling of Rural South African Children: Sibling Synergy and Migrant Effects

Katy Cornwell[a], Brett Inder[b], Pushkar Maitra[c*] and Anu Rammohan[d]

[a] Department of Econometrics and Business Statistics, Monash University, Clayton Campus, VIC 3800, Australia. E-mail: Katy.Cornwell@Buseco.monash.edu.au
[b] Department of Econometrics and Business Statistics, Monash University, Clayton Campus, VIC 3800, Australia. E-mail: Brett.Inder@Buseco.monash.edu.au
[c] Department of Economics, Monash University, Clayton Campus, VIC 3800, Australia. E-mail: Pushkar.Maitra@Buseco.monash.edu.au.
[d] Anu Rammohan, School of Economics and Politics, University of Sydney, Sydney, NSW 2006. Australia. E-mail: a.rammohan@econ.usyd.edu.au

Acknowledgements

Funding provided by Faculty Research Grant Scheme, Faculty of Business and Economics, Monash University. We have benefited from discussions with Ana Deumert, Gary Fields, Lata Gangadharan, Louise Grogan, Siddhartha Mitra, Joe Hirschberg, Jenny Williams, seminar participants at Monash University and the University of Melbourne and participants at the Inaugural Australian Development Economics Workshop at the University of New South Wales and the Conference of Economists at University of Melbourne. The usual caveat applies.

[*] Corresponding author. Email: Pushkar.Maitra@Buseco.monash.edu.au. Phone: +61 3 9905 5832. Fax: +61 3 9905 5476.

Household Composition and Schooling of Rural South African Children:
Sibling Synergy and Migrant Effects

Abstract

In this paper we examine the demand for education among rural Black households in South Africa using nationally representative data from the 1990s. In particular our study focuses on factors affecting schooling decisions at the household level. Our estimation results reveal strong evidence of a sibling synergy effect, in that the presence of other school-age children in a household makes it more likely that a child will attend school. We also find that having working-age migrant adults improves educational participation and attainment of children. Our results point to strong gender effects, with the presence of female migrants increasing the likelihood of girls getting more education. Finally, our results show that pensions in the hands of the grandmother increases the probability of girls attending school, but has little effect on the schooling of boys.

JEL Classification: O12, I21, C25

Keywords: Household composition, Schooling, Education Attainment, Sibling Synergy, Migrant Effects, South Africa

1. Introduction

Investment in human capital through education is universally recognised as an essential component of economic development. While education endows individuals with the means to enhance their skills, knowledge, health and productivity, it also enhances the economy's ability to develop and adopt new technology. Hence, increasing education levels is an important policy concern in most countries.

Since the household is typically the decision making unit with regard to schooling decisions of children, the household decision model is the most widely used method of analysis. Hence, decisions on whether or not to send a child to school are typically made by weighing the costs and benefits of schooling. For poor resource-constrained households particularly, the cost of sending children to school including school fees, school uniforms, books, transport etc can be quite significant. Furthermore, there is also an opportunity cost attached to schooling attendance in the form of foregone labour contributions. These costs can be very high for poor households residing in rural areas, particularly where children contribute substantially both by working in the market for a wage or by contributing to household chores such as caring for younger siblings, fetching water, firewood, cleaning and so on.

There are of course economic benefits from sending children to school, primarily in the form of a wage premium – the expected rise in earnings potential as a result of greater educational attainment. Further, in developing countries in the absence of well-developed financial and insurance markets, children also act as an informal source of old-age security for their elderly parents (see Nugent, 1985, Nugent & Gillaspy, 1983). The higher the education level attained by the child, the greater is their earnings potential as adults, thus increasing the likelihood of higher potential transfers from children in their parent's retirement. At the household level, this provides parents with a motive for investment in their children's schooling. Hence the decisions on fertility and children's schooling can be regarded as being made jointly. This strand of literature has been examined in early studies such as Becker & Lewis (1973), who examine this issue in the context of a household production model where fertility and child schooling decisions are jointly made. The essential idea here is that as households are resource constrained, there is a negative relationship between family size and schooling enrolment in the household. This has been termed in the literature as the "quantity-quality" trade off. This approach has primarily been used to identify policies that are likely to result in reductions in fertility and increases in educational investment at the household level.[1]

There is another strand of literature that has focussed only on educational outcomes, treating fertility as an exogenous variable. See for example studies by Gertler & Glewwe (1992), Butcher & Case (1994), Glewwe & Jacoby (1994), Jacoby (1994), Kaestner (1996) and Psacharopoulos (1997). This approach has often been used to examine policies that might be used to maximise the effectiveness of public spending on education, including issues relating to the ability and willingness of households to bear some of the costs of provision of educational services.

In this paper we use the second approach – we examine school enrolment and education attainment holding the number of siblings exogenous and estimating reduced form regressions of demand for education. This of course does not imply that

[1] There are a number of studies using data from developed countries (US particularly) that find that children with fewer siblings obtain more schooling than those with more siblings and this negative relationship persists even when family socioeconomic characteristics are controlled for. Evidence from developing countries is however quite mixed. Our results do not support the quantity-quality trade-off in schooling and we explain our results in the form of existence of economies of scale in educating

the number and composition of siblings is not important: indeed as we see later they have very significant effects on school attendance and educational attainment.

This paper focuses on the demand for education in South Africa. In particular, we are interested in the impact of the household's composition and its resource constraints on school engagement. We appeal to a large set of South African data in search of empirical support for views put forth in the literature. In particular our analysis seeks to test a number of key hypotheses. For example, does the data support the view that older children, particularly girls, need to withdraw from education in order to care for pre-school household members? Does being a member of a household with more school-age children make it more difficult to go to school as the household resource constraint becomes binding? Is there evidence that the death of one or more parent has a detrimental effect on participation in schooling? What is the impact of working-age adults migrating from rural households in search of work? What role does the aged social pension have in easing the household's resource constraint on schooling?

In its historical uniqueness, South Africa is an extremely interesting country in which to study these issues. A racially segregated education system was possibly the central pillar propping up the apartheid regime.[2] The Bantu Education Act of 1953 centralised control of Black education and linked tax receipts from the Blacks to public expenditure on education for the Blacks. This obviously led to extreme disparities in educational expenditures – for example in 1975, expenditure on an

---

children and in the development of schooling norms within the household. We discuss these issues later.

[2] During the apartheid regime, all South Africans were classified into one of the following four races: Black (African), Coloured (Mixed Race), Indian (Asian) and White (Caucasian). We stick to this terminology to remain consistent with the data and the existing literature. Our use of this terminology does not in any way indicate support for this kind of racial classification.

average White child was nearly fifteen times the expenditure on an average Black child.[3] With the Soweto Riots in 1976 and the boycotting of schools over the 1970's and 1980's, the situation improved somewhat and more resources were allocated to schools designated for Black South Africans. However, the disparities still continued to be fairly large. Additionally, as a result of the official policies implemented by the apartheid era South African government, Black families were assigned to "homelands" based on their language, irrespective of where the household had previously resided.[4] Following the "Black Homeland Citizenship Act" of 1970, the South African government forced millions of Blacks to these "homelands" and every conceivable effort was made to restrict movement between the homelands and the Union of South Africa.[5] There were also restrictions on job eligibility and in particular Blacks could not be employed as skilled workers. The returns to education for Blacks were therefore also quite low. However, with the dismantling of apartheid, educational investment appears to be an important means by which the majority of Black South Africans can escape the vicious cycle of unemployment and poverty.

This paper analyses the questions around engagement with schooling using data from the period immediately prior to and after the official end of the apartheid regime. We examine the individual and household level characteristics that affect current schooling enrolment among rural Black South African children, using different measures of educational attainment and the years of completed education. The paper starts, however, with a theoretical section (Section 2), that develops a

[3] See Case and Deaton (1999), Thomas (1996).

[4] The homelands were designated residential regions for the Blacks during the apartheid regime. These were autonomous states within South Africa.

[5] Blacks were resident of these autonomous homelands and not of the Republic of South Africa and they required special passes to justify their presence in the big cities in South Africa.

reduced form schooling participation equation using a model of household behaviour. Data and estimation issues are discussed in Section 3, followed by the empirical results in Section 4. Conclusions and policy implications are discussed in the final section.

2. The model

The theoretical model considered below is used to motivate the empirical analysis. Consider a household comprising of parents with both male (*m*) and female (*f*) children. The household maximises utility over two periods subject to budget constraints operating in each period. Assume that consumption is the only direct source of utility. In period 1, parents make decisions on how many children to have and the proportion of child hours to allocate to wage labour and to schooling. They also care for their elderly parents by making transfers. Second period consumption of the parents depends on the transfers that they receive from their adult children. Assuming there is no borrowing from capital markets to finance schooling, the budget constraints in periods 1 and 2 are:

$$c_1 = y\left(1-\tau_1\right) + n_{ui}w - n_{si}h - f\left(n_{ui} + n_{si}, \bar{n}\right) + \Omega_0 \quad i = m, f \tag{1}$$

$$c_2 = g\left(n_{si}, \bar{n}_s\right)\tau_2 \tag{2}$$

where c_1 and c_2 are parent's consumption in periods 1 and 2 respectively. The term y refers to total household income (assumed exogenous) and the proportion of income that is transferred to old parents is given by τ_1. We assume that these transfers are exogenously determined by social norms.

Let *n* be the total child hours, n_{ui} refers to child-hours spent in the labour force and n_{si} is child hours spent in schooling $(n = n_{ui} + n_{si})$. There is no leisure in this model, so total child hours correspond to child numbers so if the length of childhood is constant, child labour hours correspond to number of children in the labour force and child schooling hours correspond to the number of children in school. In addition we assume that $n_u \geq 0, n_s \geq 0$ and that households are constrained by a biological maximum number of children, $\bar{n}$, with $0 \leq n_u + n_s \leq \bar{n}$.

Child schooling costs (for example school fees, lunch expenses, transportation costs etc) are given by h. These costs are the basic subsistence costs incurred on all children and are assumed to be independent of schooling costs. Child rearing costs, such as the opportunity cost of parental time, expenditures on food, education and medicine are typically low in developing countries.

However, we further assume that there are positive "spill-over effects" from having other school-age children in the household. These spill-over effects can take two forms. First, there are economies of scale in child costs, both in terms of child rearing and schooling costs. This can come through sharing of clothing, books, parental time etc.[6] Second, having an older sibling already in school increases the likelihood of a child wanting to be in school. For example, having a child already in school may increase the social expectation on the household to also send subsequent children to school. For a child, having a sibling in school already may increase the attractiveness of school through peer group pressure. The household internalises these

[6] In this case having an additional child will lower the average and marginal cost of school attendance of the elder siblings as long as the cost of transferring these goods to children from other families is more costly than transferring them to children within the same household, an assumption that is likely to hold.

spill-over effects when making decisions on the optimum number of children to have and the amount to educate each child. One way of thinking about this is that the parent (or the household head) is the social planner or benevolent dictator within the household who is able to internalize this spill over effect. This implies that $f'(n) > 0$ and $f''(n) < 0$ for those households where there is more than one child in school.[7] Alternatively, if there are no 'spill-over effects' or if there are no other school-going children in the household, $f'(n) > 0$ and $f''(n) > 0$.

In the second period, parent's consumption depends on the transfers that they receive from their children $\left(\tau_2 g\left(n_s, \overline{n}_s\right)\right)$. The term $g(.,.)$ is the earnings function, which depends on the child-hours spent in schooling and the spill over effect from having other children in school. A fraction of this income $\left(\tau_2\right)$ is transferred by the children to their elderly parents. So the old age transfer received by parents (from a particular child) depends not only on the educational attainment of this particular child but also on the educational attainment of their other children. This essentially implies that the quantity-quality trade-off or the sibling competition effect, that has formed the basis of much of the literature on parental investment in children, might not be an obvious outcome. Indeed later we show that in the context of South Africa there is evidence of "sibling synergy" such that the greater the number of children in the school-going age group, the greater is the probability of school enrolment of a particular child.

[7] We note that this assumption raises the likelihood of a corner solution where the household has many children and sends them all to school. In this case, the household will have no income from child labour.

To avoid the problem of zero income in period 2, we introduce a means tested old-age public pension scheme. Let c_2^* refer to critical minimum old-age consumption. We can relate the means tested public scheme to c_2^* as follows. The amount of pension received by the household is given by

$$\Omega = \text{Max}\left\{0, c* - \tau_2 g\left(n_s, \overline{n}_s\right)\right\} \tag{3}$$

The second-period budget constraint can now be re-written as:

$$c_2 = \text{Max}\left\{c_2^*, \tau_2 g\left(n_s, \overline{n}_s\right)\right\} \tag{4}$$

where equation (4) defines the relationship between means tested public pension and second period consumption. Accordingly households that do not receive old-age transfers from their educated children $\left(\tau_2 g\left(n_s, \overline{n}_s\right)\right)$ are eligible for a means tested public pension. This ensures a minimum second period consumption of c_2^*. Differentiating c_2 with respect to n_s we get

$$\frac{\partial c_2}{\partial n_s} = \begin{cases} 0, \text{ if } \Omega > 0 \\ \tau_2 g'(.,.), \text{ otherwise} \end{cases} \tag{5}$$

Thus, sending an additional child to school in period 1 yields $\tau_2 g'(.,.)$ in period 2, and the household does not receive old-age public pension.

Assuming that consumption is the only source of utility, the household's problem then is to maximise utility of consumption in period 1 and 2 subject to the budget constraints given by (1) and (4). The problem is to choose n_u and n_s to:

$$\text{Max } u(c_1, c_2) = u(c_1) + \beta u(c_2) \tag{6}$$

The discount factor is given by β where $0 < \beta < 1$. The utility function is assumed to be twice differentiable with $u' > 0$ and $u'' \leq 0$. Substituting (1) and (3) into (6), the household's optimisation problem becomes:

$$\text{Max } u(y(1-\tau_1) + n_u w - n_s h - f(n_u + n_s, \overline{n}_s) + \Omega_0) + \beta u(g(n_s, \overline{n}_s)\tau_2 + \Omega) \tag{7}$$

The first-order conditions for an interior solution are:

$$\begin{aligned} \frac{\partial u}{\partial n_u} = u'(c_1)(w) - u'(c_1) f'(.,.) &= 0 \text{ if } \overline{n} > n_u > 0 \\ &\leq 0 \text{ if } n_u = 0 \\ &\geq 0 \text{ if } n_u = \overline{n} \end{aligned} \tag{8}$$

$$\begin{aligned} \frac{\partial u}{\partial n_s} = \beta g'(.,.)\tau_2 u'(c_2) - u'(c_1)(h) + u'(c_1) f'(.,.) &= 0 \text{ if } \overline{n} > n_s > 0 \\ &\leq 0 \text{ if } n_s = 0 \\ &\geq 0 \text{ if } n_s = \overline{n} \end{aligned} \tag{9}$$

From (9), we have the tangency condition,

$$\frac{u'(c_2)}{u'(c_1)} = \frac{h + f'(.,.)}{\beta g'(.,.)\tau_2} \tag{10}$$

Equation (10) is the arbitrage condition for shifting consumption between the two periods through investment in children's schooling. An interior solution is characterised by the tangency between the "full prices" and the marginal rate of substitution between the two periods.

In the analysis that follows, we will examine empirically some of the key features of this theoretical model. Specifically, we will look for evidence of these "spill-over effects", which suggest that children with siblings may be more likely to engage in schooling. We will also look at the effect on schooling decisions on the migration of working-age adults to the urban areas, as well as considering the role of the old-age pension. Our empirical model can be thought of as a reduced form demand equation for education, drawing on the features highlighted in the theory.

3. Data, descriptive statistics and estimation issues

We use several different data sets for our econometric analysis. The first is the South Africa Integrated Household Survey (SIHS) collected in 1993 as part of the World Bank's Living Standard Measurement Surveys undertaken in a large number of developing countries. This data was collected in the nine months prior to the first democratic elections that brought Nelson Mandela to power. The survey was conducted jointly by the World Bank and the South Africa Labour and Development Research Unit (SALDRU) at the University of Cape Town. This cross-sectional data set is unique because it is the first survey that covers the entire South African population, including those in the predominantly Black "homelands". The sample consists of approximately 9,000 households drawn randomly from 360 clusters. The questionnaire and summary statistics are contained in SALDRU (1994). We will henceforth refer to this data as the SIHS1993 data set.

Next we use the October Household Survey (OHS) data sets from 1996 – 1998. The OHS is an annual survey, based on a probability sample of a large number of households. The data sets cover a range of development indicators that include demographic factors (such as age, gender, level of education, marital status, migration, use of health services, internal migration), economic variables (such as employment, unemployment, informal sector employment) and social and well being measures (access to health and social services, safety of household, average household size, type of dwelling, level of education, quality of life, health statistics, vital statistics). Though the OHS's have been conducted every year from 1993 onwards, our choice of years is dictated primarily by data availability and data comparability across the different surveys. Each of the OHS's are essentially independent cross-sectional surveys.

Our primary objective is to identify the determinants of school attendance and educational attainment. Individual children are the unit of observation, and we focus on the following two dependent variables: (1) Current enrolment (ATTSCH): a dummy variable taking the value 1 if the child is currently enrolled, and 0 otherwise; (2) Grade attainment (ATTAINED): the highest grade achieved by the child. Both these variables provide a good estimate of the extent to which children are combining schooling and work.

Current enrolment is analysed separately for children in the age groups 7 – 12 and 13 – 18, as it is expected that different sets of explanatory variables affect the decision to attend school for the two age groups.[8] When estimating grade attainment, several important estimation issues arise. Because of the possibility that children may start schooling late, the age groups and education level attained do not necessarily match perfectly and therefore the final complete level of education for the members of the cohort is not necessarily known at the time of the survey. Those individuals who are still in school can hence be regarded as being censored – a common problem present in cross-section data of this kind. We therefore restrict the sample to children aged 13 – 18, as children in this age group should normally have completed primary schooling. This will eliminate some but not all of the censoring effect. To the extent that this censoring is related to the individual, household level and community characteristics that are used as explanatory variables, this may bias the estimated coefficients.[9]

A further issue associated with modelling grade attainment centres on potential selection bias created by the fact that some in the sample were never enrolled in school. OLS estimates of grade attainment are thus vulnerable to selection bias. We could consider a Heckman selectivity-correction model for grade attainment conditional on enrolment. This approach was rejected because of the absence of credible exclusion restrictions (identifying variables that affect the probability of enrolment but not grade attainment). As an alternative, we decided to retain the never-enrolled children and use an ordered probit model, where grade attainment is re-organised into three hierarchical categories. The categorical grade-attainment variable (ATTAINED) takes the value 0 for never-enrolled children, 1 for ever-enrolled children who have not completed the primary stage (i.e. five years of schooling), and 2 for those who have completed the primary stage. In this case the sample is restricted to children aged 13 – 18, who under normal progression rates, would have completed primary schooling.

For purposes of estimation the sample is restricted to Black households residing in rural areas of South Africa. This overcomes the complexities of historical race-related variable interactions, and refines the discussion on migration to be

---

[8] We also estimated current enrolment status for all children (aged 7 – 18) but decided not to present these results. They are available on request.

[9] When current school enrolment is the dependent variable, this censoring issue is not a problem. However, there might be other problems associated with this particular variable. The estimation results will be biased if either the age at which children start schooling or the extent to which they repeat grades is related to the explanatory variables that are used. For example, suppose that girls and boys have identical education attainment but girls start school later or repeat more grades. In this case, at any point in time we expect to see more girls than boys enrolled in school, giving the false impression that girls receive more education than boys. For more on these issues see Anh, Knodel, Lam &

primarily rural-outgoing migration. A list and definition of all variables used in the models is presented in the Appendix, and Table 1 presents descriptive statistics for a selection of the variables. The explanatory variables that we use in the regressions include individual, household level and community level characteristics.

Individual (child level) characteristics include the age (AGE), age-squared (AGESQ) and child's gender (MALE). Previous research from Africa (Garg & Morduch, 1998, and Morduch, 2000) have argued that sibling gender composition has an important influence on intrahousehold resource allocation of schooling and health resources, particularly if the child comes from a poor, resource-constrained household. For example, if there is a gender bias operating at the household level, then for a given family size, it must be the case that a male child, growing up in a household with only brothers, may have fewer resources than if he were to grow up with sisters only. This is likely to be true for females as well. This would imply that the educational attainment of children depends not only on their own gender but also on whether their siblings are male or female. Hence, siblings become rivals in a competition for greater access to household resources. To take into account this effect, the number, age and gender of siblings is stratified as follows: number of male and female siblings in the age group 0 – 6 (SIBM_0-6 and SIBF_0-6), number of male and female siblings in the age group 7 – 12 (SIBM_7-12 and SIBF_7-12) and number of male and female siblings in the age group 13 – 18 (SIBM_13-18 and SIBF_13-18). To examine whether sibling composition has differential effects on

Friedman (1998). While we acknowledge these issues, it should be noted that the existing data does not allow us to correct for these potential problems.

schooling of boys relative to girls, we also interact each of the sibling composition variables with the gender of the child (MALE) dummy.

While sibling composition is likely to play an important role in a child's schooling and educational attainment, it is not very clear *a priori* how it actually does so. Indeed the evidence on the effect of sibling composition on educational attainment is mixed. Parish & Willis (1993), using data from Taiwan, find that having sisters can be an important predictor of schooling outcomes. In particular they find that having older sisters is especially beneficial for younger siblings as these older sisters are not only more likely to help take care of their younger siblings, they are also more likely to take up wage employment that helps pay school fees and allows younger children to postpone entering the labour market. Kaestner (1996) finds that in a sample of African-American men in the US, completed schooling would increase by 0.77 years if they had all sisters as opposed to all brothers. In the context of Africa, Morduch (2000) finds that in Tanzania moving from a situation where a particular child has all brothers to one where he/she has all sisters, the number of completed years of schooling increases by 0.44 years. He does not however find any inherent gender effects. In South Africa however Morduch (2000), finds no statistically significant gender composition effect.

Turning to adult characteristics, we include three dummy variables to indicate the mortality status of the parents: mother deceased (MOTHERDEC), father deceased (FATHERDEC) and both parents deceased (PARDEC).[10] There is now a fairly large literature that examines whether child outcomes are different depending on whether

[10] Note that the way the data is coded, when PARDEC = 1 (i.e., both parents are dead, MOTHERDEC = 1 and FATHERDEC = 1. So to obtain the "full" effect of both parents dead, we need to add the coefficient estimates of MOTHERDEC, FATHERDEC and PARDEC.

the child lives with their parents or not. In Sub-Saharan Africa in particular, this is related to the widespread practise of child fostering (see for example Akresh, 2004). In South Africa during the 1990's this issue has taken a rather unexpected twist given the spread of the AIDS pandemic in the rural areas. Hence a child whose parents have died (either because of AIDS or otherwise) might be more likely to work, might experience psychological problems, or might suffer due to the disruption of having to live away from his or her siblings.[11]

We include a number of household control characteristics. The first is log per capita household expenditure (PCEXP), used as a proxy for household permanent income. Household expenditure is easier to measure compared to household income and is typically measured with less error and is also considered a better proxy for permanent income because while income might be subject to transitory fluctuations, households use a variety of mechanisms to smooth consumption over time.

The second household characteristic of interest is to do with the presence of migrant workers. Since migration in most cases is a response to potential employment opportunities, having a migrant household member can imply greater transfers to the household, hence easing the household's resource constraints. A perusal of the age composition of Black households residing in rural South Africa reveals a rather interesting pattern. Households are multi-generational but are often characterised by a missing middle generation: working age adults (males and females) leave their children with their grandparents and migrate to the cities in search of jobs. Figure 1 shows quite dramatically how rural areas are over-represented with children and the elderly, and grossly under-represented in the working-age population. This form of

[11] We would have liked to include parents' educational attainment as additional explanatory variables,

migration has a long history in South Africa. For many years Black households in rural South Africa have been dependent on income transfers from members of the household living away and working in the mines and plantations.[12] After the dismantling of apartheid and the repealing of the "pass laws", Black women now have the legal freedom to migrate to the cities in search of employment. However because of the poor living conditions in South African cities, adults often prefer not to bring their children with them to the urban areas, instead leaving them with their grandparents and/or other relatives in the villages.

It is difficult to predict the effect of the presence of a migrant in the household on child schooling. On the one hand, the absence of working-age adults could mean a greater likelihood that children might have to contribute by working around the house, thus having an adverse effect on schooling. This could be particularly important for the older children (those aged 13 – 18).[13] On the other hand, to the extent that migrants remit a portion of their income back home, this could help ease the household's resource constraint and may increase the likelihood of a child's participation in schooling. Moreover, if these migrant workers are perceived as prosperous (comparatively), this could act as an incentive for the rural household to send their children to school: on the assumption that an investment in human capital stock enables the child to find a better job in the city on completion of their education. We term this second effect the "migrant effect".[14] We examine the effect of migrants by including dummies for whether a male and a female member of the

but this was not possible because the data is not available in these surveys.

[12] See Wilson (1972) for a history of the migrant labour system in South Africa.

[13] See Lipton (1980) for a discussion.

[14] See for example Stark, Helmenstein & Prskawetz (1997). This line of research has been broadly categorised as the new economics of labour migration.

household is absent as migrant workers (MIGRANTM and MIGRANTF respectively). We also interact these dummies with the gender of the child to examine whether there is a differential impact of the presence of a migrant in the household on schooling and educational attainment of boys and girls.

The final set of variables centre around whether the household includes an aged pension recipient. When in 1993 the aged pension was extended to non-White elderly, it represented a significant boost to the disposable income of rural black households. In 1993 the maximum benefit of R370 per month was paid to all women above the age of 60 and all men above the age of 65, subject to a "means" test.[15] By 1999 the monthly pension payment had risen to R520 per month. It has been argued that increased access to the social pension scheme may well have encouraged households to have children and other dependents live with the pensioners. Edmonds, Mammen & Miller (2003) use census data from South Africa for 1991 and 1996 to support this view, as do Bertrand, Mullainathan & Miller (2003) using the SIHS1993 data.[16]

In terms of the impact on schooling, our hypothesis is that access to social pension has two possibilities: first it eases the resource constraint on the households, hence increasing the probability of schooling for children. Second, the availability of pensions also reduces the need for old-age security from children, which removes an

[15] For a single age-qualified individual the "means" is defined as the sum of income and the imputed value of income-generating assets. In 1993, the pension would be reduced one-for-one when the means exceed R 90 a month until the means reach R 370 a month beyond which pension payment stops. For age-qualified couples, the means is calculated by pooling and splitting equally the income and imputed value of income-generating assets. Note that the means test does not take into account the income of the other members of the household. For the Black South African households the "means" is set at such a high level that it is not binding for most households. See Alderman (1999).
[16] Edmonds, Mammen & Miller (2003) find that households that receive social pension have fewer working prime age adult females but have more resident children below the age of 5 and young women of childbearing age. Bertrand, Mullainathan & Miller (2003) on the other hand find that access to old

important motivation for schooling investments. However, since we assume that there are spill-over effects from schooling, schooling costs are considerably reduced, hence it is more likely that an easing of resource constraints increases the probability of schooling. Previous research by Duflo (2003) has also found that separating pension receipt by gender has different affects on child health investments. Hence, we also separate by the gender of the pension recipient since it allows us to examine whether the child's schooling differs depending on whether the male or the female household member receives the pension.

An appropriate explanatory variable for pension receipt would be the actual amount of pension received by members of the household. However, for several of the surveys, this data seems to contain a number of errors – often the amount of pension received takes quite implausible values. Consequently, we decided to use the more indirect yet more reliably measured variable capturing the number of adults who are age-qualified to receive the pension (males aged 65 and over, females aged 60 and over). This variable should actually provide a good indication of amount received: in most cases the pension is a fixed amount per eligible recipient. The income means test, which could lead to a reduced pension receipt, applies to very few Black rural households.

Distance to the nearest school is typically used as the cost of schooling, as other measures of the cost of schooling, like fees paid, costs of books, costs of travelling etc could be correlated with the unmeasured determinants of the demand for schooling, leading to standard endogeneity problems. In the absence of data on distance to the nearest school, for models using the SIHS1993 data we include the

---

age pension by members of the household results in a sort of moral hazard problem in that it increased number of primary schools (NUMPS) and the number of secondary schools (NUMSS) in the particular community as the relevant supply-side variables. Unfortunately, there is no equivalent supply-side data available for the OHS. We tried to capture these supply side variables by two alternative methods. First we included the distance of the house where the child resides from the nearest source of water. In rural South Africa typically, it is the responsibility of the child to fetch water from the nearest source, and in some sense we could think of this distance as a component of the cost of schooling.[17] Second, we estimated a random effects probit/ordered probit model to capture the (unobserved) cluster effects, which could be viewed as proxies for the supply side variables. However here we are restricted to using data from the 1996 and 1997 surveys since for 1998 survey the cluster/sample enumeration area is not adequately defined.[18] However in either case, the main results (both qualitatively and quantitatively) remain unaffected.

The indications are that if we were not constrained by data availability and were in fact able to include these supply side variables, it is likely that the results would not change. Accordingly in the results that we present in this version of the paper, we decided to not include the distance to water variables and also decided not to take into account these unobserved cluster effects.[19] We include a set of province dummies to account for any other unobserved heterogeneity. Notably, it was not until 1995 that South Africa's 14 provinces and homelands were restructured into the current 9 provinces. Hence, in order to capture municipal boundary heterogeneity, we

---

the probability of unemployment of male working age adult members of the household.

[17] Of course we should remember that this variable could be endogenous in that it could be correlated with the unobserved determinants of school enrolment and grade attainment.

[18] We could not compute the probit estimates with the cluster fixed effects because of convergence problems: we have more than 1700 clusters.

[19] These results are of course available on request.

use different province dummies in our analysis of the SIHS and OHS datasets. The reference category is that the household resides in Cape (SIHS) and Western Cape (OHS) provinces.

Another methodological issue to consider in the estimation is the possibility of endogeneity of household permanent income/expenditure in the demand for schooling equation – household income is likely to be correlated with the unmeasured household-specific determinants of child schooling. The instruments used to overcome this endogeneity include the demographic and educational characteristics of the household head and the characteristics of the household where the child lives. We use the two-stage methodology developed by Rivers & Vuong (1988) to correct for this potential endogeneity problem.[20]

4. Discussion of results

Table 2 presents the OLS estimates for the first stage regressions used to obtain PCEXPE. As mentioned previously, separate estimates are provided for the SIHS 1993 data and the OHS 1996 – 1998 data. While this equation is not of primary interest (being used essentially to deal with endogeneity in the schooling equations), some results are worthy of comment. Not surprisingly our results suggest that the higher the education level of the household head, the greater the permanent income of the household. Only in the SIHS1993 data is there evidence that male-headed households are economically better off compared to female headed households, while

[20] In the first stage, PCEXP is regressed on the set of instruments. The errors from this regression are then included as an additional regressor in the second stage probit or ordered probit estimation. A significant coefficient on PCEXPE implies rejection of the null hypothesis of exogeneity of permanent income. This null hypothesis is always rejected in this study; estimates would be inconsistent if we did not control for this endogeneity.

age of the household head appears to have a clear relationship for both data sets. Several of the household infrastructure and living conditions variables also turn out to be important, all taking the expected signs.

We now turn to the estimation results regarding school attendance and attainment. There are quite a number of results to discuss, so to make the discussion manageable, we discuss the results for two particular specifications in detail, and then highlight ways in which the alternative specifications either confirm or vary from these.

In each regression we include all the children in the sample and as a result some households contribute multiple children to the sample. Since unmeasured determinants of schooling are likely to be correlated between families (or households), the estimated standard errors will be biased downwards. We allow for this correlation by adjusting the standard errors for clustering on households.

4.1 Current school enrolment

Table 3 presents the first set of results where the dependent variable represents whether the child is currently in school. Discussion will focus mainly on the models using OHS data sets (presented in Columns (3) and (4)) since these involve a larger number of observations, and cover a period of relatively greater political stability, making them more likely to be indicative of the underlying relationships. Rather than give the estimated probit coefficients, the table presents the more easily interpreted marginal effects. These are defined as partial derivatives of the probability of a child being enrolled in school with respect to the individual control variables, holding all dummy variables to zero and all other variables at the sample means. As noted earlier, we compute and present separate estimates for school-age children in the two age categories: the 7 – 12 and 13 – 18 year olds.

Consider first the estimated model for 7 – 12 year-old children in the OHS data set. The gender dummy is not statistically significant implying that on an average both boys and girls are equally likely to be enrolled in school. However, several of the explanatory variables have different impacts on the schooling of boys and girls, hence we cannot simply use this result to conclude that there are no gender differences in schooling. There is an inverted u-shaped relationship between the age of the child and the probability of school attendance: an increase in the age of the child increases the probability that the child is enrolled in school but the probability falls after reaching a maximum (attained at around 10.3 years of age). Notably, none of the sibling effects are significant, indicating that there is no evidence that a 7 – 12 year-old's attendance at school is affected by the presence or otherwise of other children in the household. This finding contrasts with results for older children, which will be discussed below.

Turning to the characteristics of the adults in the household, a result that initially seems counter-intuitive is that where the mother or the father is deceased, the child is actually more likely to be attending school. The effect when both parents are deceased is found by adding together the coefficients on the three variables MOTHERDEC, FATHERDEC, and PARDEC. Not surprisingly, this produces a strong negative marginal effect of –0.077 implying that being orphaned has a detrimental effect on a child's schooling. What is puzzling however is that when only one parent is deceased, this actually improves the child's school attendance. This rather surprising finding can perhaps be explained by the possibility that the surviving parent now feels a greater need to educate their child in light of the loss of one parent.

Over the period 1996 – 1998, 16% of households have at least one female adult migrant and 26% have at least one male migrant. Figure 1 highlights the magnitude of the effect of migration on the age composition of residents in the rural areas relative to urban areas. A ratio of one indicates that in the given age group, the proportion of the population in this age group is the same in both rural and urban areas. The U-shaped curves demonstrate how the rural areas have disproportionately high representation of children and elderly, and relatively low shares of working-age adults. For example, in 1996 – 1998, the ratios for 25 – 44 year-olds are all around 60%, indicating a significantly smaller representation of those at prime working-age in the rural areas. These differences provide us with a motivation for a closer look at the effects of migration on schooling. The estimated coefficients for MIGRANTF and MIGRANTM suggest that when a household has a male and/or a female migrant, this actually produces better school attendance among 7 – 12 year-olds. The gender interaction terms (M_MIGF and M_MIG_M), however, have negative coefficients, though not statistically significant. Taking into account these offsetting negative effects, we find that the effect of the household sending a migrant on boys is very small. The estimation results suggest that for girls, the “migrant effect” (households with migrants are better able to recognise the greater returns to education and hence children from households that have migrants are more likely to attend school) dominates, while the two effects tend to cancel out for boys.

The next set of coefficients centre around the presence of elderly residents in the household. Neither variable measuring the number of adults who age-qualify for the aged pension (AGEQF and AGEQM) seems to have any effect on attendance of these young children in schooling. This result may appear surprising, given that the aged pension provides a relatively large source of income to poor Black households, and we should expect to see greater schooling attendance. However, note that the equation also includes the household expenditure variable, which by providing some measure of income/wealth of the household, captures the main resource constraint effect. In other words, consider two households with the same level of household expenditure, with one having no age-qualified adults, and the other having one. The coefficient on the relevant age-qualified variable measures the average differential effect on schooling participation of being part of the household with the aged-qualified person. Once we have controlled for expenditure, there is probably no strong reason to believe that pension receipt or the presence of elderly household members would add to or take away from schooling participation.

The coefficient for household expenditure variable (PCEXP) shows a strong positive effect, indicating that a resource constraint on schooling for 7 – 12 year-old children is very real. The marginal effects imply that a 100 Rand increase in monthly per capita household expenditure/permanent income leads to a 3 percentage point increase in the probability of school enrolment for the 7 – 12 year old children.

Finally, the 1997 and 1998 year dummy variables both have significant negative coefficients, indicating a deterioration in schooling attendance, after controlling for other factors - at quite a high rate of at least 3.5% per year. This is consistent with the pattern in the raw data, as presented in Table 1.[21]

[21] Marginal effects of the Provincial dummies are not reported in the tables. These are generally quite significant.

There are some notable differences in the results for the 13 – 18 year-old children using the 1996 – 1998 OHS data sets (Column (4) of Table 3). The direct measure of gender effects is also not significant, and the role of the child's age is qualitatively the same. However, sibling effects are quite different. Although no sibling effects were found with 7 – 12 year-old children, with these older children there are two dominant sibling effects. First, when the household includes 0 – 6 year-old children (pre-school age children) (SIBF_0-6 and SIBM_0-6), this has a strong negative effect on school attendance. However, the gender interaction terms (M_SIBF_0-6 and M_SIBM_0-6) largely cancel these negative effects, suggesting that 13 – 18 year-old girls from households that have young pre-school children have poorer school attendance, but the presence of these pre-schoolers has little or no effect on the school attendance of 13 – 18 year-old boys. This is quite a plausible result, suggesting the traditional roles where older girls are often expected to mind younger children, but the same is not expected of boys. Consequently, it is the schooling of teenage girls that suffers. Secondly, virtually all the other sibling effect variables (SIBF_7-12, SIBF_13-18, SIBM_7-12, SIBM_13-18) have positive and significant (or almost significant) coefficients. In other words, the presence of other school age children in the household tends to increase the probability of a 13 – 18 year old attending school.

The gender interaction variables are all insignificant, indicating that this result applies to both girls and boys. This finding is interesting, and in contrast to the "sibling rivalry" hypothesis that has been advanced, where households choose to direct their limited resources to schooling selectively depending on which children are likely to generate better returns to education. Instead, we find that the more school-age children there are in a household, the better the school attendance of 13 – 18 year-olds. Perhaps what we are observing here is more a "sibling synergy" effect, whereby households with several children have the synergies that give greater encouragement to these children to attend school. Somehow the resource constraint that would go along with having to educate a larger number of children does not seem to have much impact compared to these synergies. As described in our theoretical model, these synergies may be explained by the "spill-over effects" of having other school-age children in the household – taking the form of economies of scale in child costs, and the development of a culture of schooling and educational attainment within the household. Furthermore, having an older sibling already in school increases the social expectation on the household to also send their other children to school.

There are a number of studies using data from developed countries that find that children with fewer siblings obtain more schooling than those with more siblings and this negative relationship persists even when family socioeconomic characteristics are controlled for.[22] Evidence from developing countries is however quite mixed. Evidence from Thailand (Knodel, Havanon & Sittirai, 1990) and Brazil (Psacharopoulos & Arriagada, 1989) suggest that there is a negative relationship between the number of siblings and educational attainment. In the case of Vietnam (Anh, Knodel, Lam & Friedman, 1998) the relationship is negative for families with six or more children and the effects are quite small once other families are controlled for. In the context of Indonesia, Maralani (2004) finds that the relationship between sibling size and schooling was positive or neutral for the older cohorts, but for the younger cohorts there is a negative relationship. Qian (2005) uses data from China and finds that for one-child policies, an additional child significantly increases the school enrolment of the first born. Using data from Israel, Shavit & Pierce (1991) find that for the richer Jews, family size has a negative relationship with educational attainment of children, while there is a positive relationship between family size and educational attainment of children in the poorer Muslim households. In African countries however there is no negative effect and in fact the opposite is true: educational attainment has a positive relationship with the number of siblings (see Gomes, 1984 for evidence from Kenya & Chernichovsky, 1985 for evidence from Botswana). The explanation for this positive relationship typically involves households in Africa (and indeed the poor Muslim communities in Israel) drawing on a large kinship network beyond the immediate family, which reduces the costs (financial, emotional and time) associated with additional children. Our results from South Africa tend to support this argument.

The next main point of difference with the 13 – 18 year-old children is with the impact of deceased parents on schooling. With these teenagers, the death of either one or both parents has no effect on participation in schooling. This is not surprising, as 13 – 18 year-olds are less dependent on parental care and support than younger children. The effect of the household sending a female migrant is essentially the same for both age groups. However, the effect of sending of a male migrant appears to have weakened considerably for the older age group, to the extent that it now seems any effect is felt by teenage boys rather than girls. The presence of age-qualified female elderly in the household appears to improve the chances that a girl attends school, yet

[22] See for example Stafford (1987), Behrman, Pollack & Taubman (1989) and Conley (2004) for the

there is essentially no parallel effect for boys. Household expenditure has a similar effect on school attendance for both age groups.

Finally, while there was an underlying decline in school attendance observed in both 1997 and 1998 with 7 – 12 year-olds, the decline is found only in 1998 with these older children.

We now discuss other results on school attendance where the data set or specific included variables are varied. First, Columns (1) and (2) of Table 3 present results for the model discussed above using the 1993 SIHS data set. The only difference in the specification of this model is that we are now able to consider two supply-side variables (NUMPS and NUMSS) relating to the number of primary and secondary schools in the community. Both these supply-side variables turn out to be statistically insignificant.

The remainder of the results for both the 7 – 12 and 13 – 18 year-olds are largely the same, although fewer significant effects were found. This is most likely because of the much smaller sample in these estimates, and the fact that during 1993 South Africa was in a state of political transition, which would possibly cloud the underlying structural relationships, particularly for rural Black households.

One difference, however, is found in the significance of age-qualified elderly for the 13 – 18 year old age group – while the signs are mostly the same, the effects are much stronger. In the 1993 sample, it appears that having elderly household members does impact on child schooling. For example, having elderly females in the household improves the chances that a teenage girl will attend school, while the opposite appears true for the attendance of boys. The presence of elderly males in the

---

US and Goux & Maurin (2004) for France.

household on the other hand reduces the chances of teenagers attending school, regardless of whether the child is a boy or a girl.

One explanation for this is that in 1993 the pension had just been made available to all South Africans, and so at this transitional time being age-qualified for the pension would not necessarily imply receipt of the pension, particularly for rural Black South Africans. To examine the plausibility of this explanation, we consider an alternative measure of pension receipt by including the actual number of pension recipients (RECPM and RECPF). Indeed, we find that the SIHS1993 results do differ depending on whether pension receipt variable is derived on the basis of age-qualification or recorded pension receipt (Columns (1) and (2) of Table 4). In agreement with the OHS results (which are particularly robust to the alternative measure of pension receipt), pension in the hands of the grandmother improves the chances of girls attending school, but has little effect on the schooling of boys.

Not only does this result affirm that the difference in results between the SIHS1993 and OHS 1996-98 may be due to slow transfer of information, it suggests that the pension, by way of its contribution to household income, has helped to get children, and particularly girls, to school.

This gender-specific pension effect has been examined and confirmed in the literature. Using the same SIHS1993 data set, Duflo (2003) finds that an increase in female pension receipts have a positive and statistically significant effect on the anthropometric status of girls in the household. She argues that this is evidence against the standard unitary model of the household. We find a similar result in the context of child schooling.

Table 5 reports on attempts to consider more closely the effect of migration on schooling. One possible reason for a positive association between migration and schooling could simply be that migrants remit money to households thus easing their resource constraints somewhat. Alternatively, and perhaps the more likely explanation is that the migration experience generates greater motivation for schooling: upon their return home, the migrant worker brings with them experiences, ideas and physical items that are not available in the rural areas. In fact, it seems almost essential that the migrant goes to great effort in portraying the "bright lights" image of urban life. Migration is then seen as the gateway to materialism and prosperity. Investment in education is thereby a means by which to improve the chances of finding work in the urban area, and hence a "way out" of the village. If this were the case, we argue that this effect would be captured not so much by whether a migrant has been sent from within the household, but more what proportion of households in the village have sent a migrant. The models reported in Table 5 thus replace the migrant variables with a variable representing the proportion of households in the cluster that have at least one male and/or female migrant (PROPMIGM and PROPMIGF), together with gender interaction terms (M_PROPMIGM and M_PROPMIGF).[23] The OHS results suggest positive migrant effects on schooling for 7 – 12 year-old girls, but not so much for boys, and no significant effects for the 13 – 18 year-old age group. This is very similar to the results when the migrant variables represent migration from the particular household.

On a final note, it could be argued that our models have not accounted for household size. However, we note that we have implicitly accounted for this through

[23] Recall that in this case for the OHS we are restricted to using data from the 1996 and 1997 surveys.

our household composition variables (number of siblings, parental presence, sending of migrants, number of elderly etc.). In fact, including the number of household members as an additional explanatory variable not only proved insignificant, but it also did not change any of the estimated coefficients in any of our models.

4.2 Grade Attainment

The first set of ordered probit regression results for grade attainment are presented in Table 6. The dependent variable is coded to take the value zero if the child has never attended school, 1 for completion of some primary level schooling, and 2 for those who have completed primary school. Recall that the sample is restricted to children aged 13 – 18.[24] As before, we initially focus discussion on the results using the 1996 – 1998 OHS data. Column (5) presents the estimated coefficients, and Columns (6) – (8) marginal effects evaluated at sample means for continuous variables and zero for dummy variables.

The first point to note from the results in Table 5 is the significance of the dummy for child's gender (MALE). These marginal effects imply that a boy is around 10 percentage points less likely to have completed primary schooling than a girl. Contrast this with the results on school attendance, where no direct gender differences were found: this evidence suggests that boys' lower completion rates are not due to poorer attendance; there must be other gender-specific factors at work. Once again we find that age has a significant non-linear effect on the probability of primary

[24] A positive coefficient on a particular explanatory variable implies that the variable increases the probability of the child aged 13 – 18 having completed primary school, implying a lower probability of having not attained any schooling or completing only some primary schooling. Conversely, a negative coefficient implies that the variable increases the probability of not having attained any schooling, decreasing the probability of having completed primary schooling).

school completion, with the coefficients of AGE and AGESQ suggesting that older children are more likely to have completed primary school.

Sibling effects are similar but not identical to those for attendance. First, we note that as with attendance, the presence of 0 – 6 year-old children in the household has a negative effect on school attainment for girls. Secondly, presence of other 7 – 12 year-old or 13 – 18 year-old children seems to have either a weak or a non-existent effect on grade attainment. In other words, there is no evidence of sibling rivalry effects, and little evidence in support of sibling synergy. This contrasts with the results for school attendance, which suggest clear benefits to the presence of school-age children in the household. The overall picture painted by these results is that there are benefits from being in a household where other school-age children are attending school, but not with achievement of outcomes.

The coefficients of the remaining variables are generally not significant, suggesting that there is no clear evidence that the death of parents, the migration of adults from the household or the presence of age qualified elderly people in the household affect school achievement. The strongly significant PCEXP variable does suggest the presence of wealth effects/resource constraints in achieving successful outcomes in schooling.

Columns (1) – (4) of Table 6 present the results for estimates of the same model, but with the SIHS1993 data set. The results tell a very similar story to those presented in Columns (5) – (8), although there are a couple of other variables which have significant coefficients. First, we find that SIBF_7-12 has a significant positive coefficient, indicating that there is some evidence of positive sibling effects on primary school attainment with this data set. Secondly, a significant coefficient on

PARDEC indicates a clear negative effect on school attainment for those children whose parents are both deceased.

For completeness, Table 7 presents the results with the alternative measure of pension receipt (the number of pension recipients in the household rather than the number of age-qualified elderly residents). The results are robust to this alternative measure, although the coefficient on female pension receipt is a slightly stronger in favour of completion of primary schooling for girls.

Table 8 presents the estimated regression results with the alternative migration variable: instead of using the number of migrants in the household, we include a variable measuring the proportion of households in the cluster that have at least one male and/or female migrant (PROPMIGM and PROPMIGF), together with gender interaction terms. The inclusion of this alternative migration measure has virtually no impact on the signs or significance of the various other variables in the model; the story with the sibling effects, pension effects, deceased parents and wealth effects remains the same as described in Table 6. However clusters with a higher proportion of female migrants have a positive effect on primary school completion – PROPMIGF has a positive coefficient for both data sets.

Finally we also examined (i) school attendance for the household as a whole economic unit: what proportion of children in a particular age group attend school, conditional on the household having a child in that age group (using a double sided tobit model) and (ii) whether the child ever attended school (using a probit model: the dependent variable takes a value of 1 if the child has any primary schooling and 0 otherwise).[25] The tobit regression results generally corroborate our previous sibling

[25] These results are available on request.

synergy results: **a**cross all estimated models we find evidence of sibling composition effects but the effects are considerably stronger when the proportion definition includes children aged 13 – 18. Boys and girls aged 0 – 6 in the household have a detrimental effect on the proportion of children in school, while primary school aged children have a positive effect on schooling for the older age group. The number of girls and boys in the household aged 13 – 18 has a positive association with the proportion of primary school aged children attending school. There are however no sibling composition effects on the probability of ever attending school.

5. Concluding Remarks

Education and human capital accumulation are essential components of economic development. This paper attempts to identify some of the individual and household level characteristics that affect the demand for schooling in South Africa, focusing on Black households residing in rural South Africa. Among the key hypotheses we test are: does the data support the view that older children, particularly girls, withdraw from school (maybe to take care of their younger siblings)? How important are the sibling competition and household resource constraints in affecting school attendance and educational attainment? What is the effect of the death of one or more parent on child schooling? What is the impact of migration of working age adults? How does social pension received by the members of the household affect school attendance of the children?

We conclude by re-emphasizing our main findings. First we find evidence of significant sibling synergy effects on schooling of children: the regression results show that attendance of 13 – 18 year olds is enhanced when there are other household members in the school going age group and primary school completion rates are higher for girls when they have other members in the school going age group. Our results are therefore supportive of the sibling synergy effect. These results corroborate those from other countries in Africa. One explanation of these results comes from the fact that there might be positive spill over effects from having children attending school, due to economies of scale in child costs and/or from the development of a "schooling culture" within the household, influenced by social forces. This spill over essentially results in a form of positive 'externality' in children's educational attainment, which the parents will try to internalize through their schooling decisions on each subsequent child. Hence, rather than the usual sibling competition effect, the model predicts sibling synergy – it is optimal for parents to educate a greater number children. Our results are in line with the predictions of the theoretical model presented in Section 2.

Second, our results tend to support the commonly held view that the schooling and educational attainment of 13 – 18 year old girls is adversely affected by the presence of pre-school age siblings – girls in this age group are more likely to drop out of school, presumably to take care of their younger siblings.

Third, there are significant migrant effects. School attendance, particularly for girls aged 7 – 12 is higher in households with a migrant. Also, an increase in the number of female migrants in the cluster or neighbourhood increases the school attendance and educational attainment rates of 13 – 18 year old girls. These results are in line with income maximising behaviour on the part of parents. The idea here is that given household resource constraints, parents invest in resources/assets that provide a greater return over the lifetime. In many developing countries this takes the form of educating boys (at the cost of educating girls) as often the lifetime return from boys is higher, driven primarily by higher male wage rates. However, Rosenzweig & Schultz (1982), using district level data from rural India, show that parental inputs for girls is higher in regions of India where female wage rates are higher. We find a similar story here: successful female migrants in the community (either in the household or in the neighbourhood) could be viewed as a signal to parents that they can expect positive returns to education even from their daughters. This results in increased investment in schooling for girls. Additionally, it is interesting to note that the absence of working age adults in the household does not appear to result in higher child labour: rather it is associated with improved schooling attendance. This might be the result of the type of farming practices in South Africa – the small size of the plots and the less hard labour intensive crop types and farming techniques which enable older household members to work to an older age, freeing time for children to attend school.

Finally, the presence of a pension receiving woman in the household significantly increases the school attendance and attainment of 13 – 18 year old girls, but this is not paralleled with boys, nor is the presence of a pension receiving man in the household conducive to schooling outcomes. These results tend to support the argument that the identity of the income recipient matters: resources are not necessarily pooled within the household.

From a policy perspective, improving the educational attainment of children in Black households is potentially an important means of bridging the gap between the different races in South Africa. Accordingly, increasing education attainment levels of the current school age population is a particularly important issue. By focussing on the factors that affect the demand for schooling (rather than supply side factors), this paper identifies several key areas for policy makers to target in their attempt to increase educational attainment and human capital accumulation.

Appendix A
Glossary of individual level variables

Variable	Description
AGE	Age of child
AGESQ	Age of child squared
MALE	Dummy=1 if the child is male
SIBF_0-6	Number of female siblings between ages 0-5
SIBF_7-12	Number of female siblings between ages 6-12
SIBF_13-18	Number of female siblings between ages 13-18
SIBM_0-6	Number of male siblings between ages 0-5
SIBM_7-12	Number of male siblings between ages 6-12
SIBM_13-18	Number of male siblings between ages 13-18
M_SIBF_0-6	Interaction of MALE x SIBF_0-5
M_SIBF_7-12	Interaction of MALE x SIBF_6-12
M_SIBF_13-18	Interaction of MALE x SIBF_13-18
M_SIBM_0-6	Interaction of MALE x SIBM_0-5
M_SIBM_7-12	Interaction of MALE x SIBM_6-12
M_SIBM_13-18	Interaction of MALE x SIBM_13-18
M_RECPF	Interaction of MALE x PENRF
M_RECPM	Interaction of MALE x PENRM
M_AGEQF	Interaction of MALE x AGEQF
M_AGEQM	Interaction of MALE x AGEQM
MOTHERDEC	Dummy=1 if the child's mother has deceased
FATHERDEC	Dummy=1 if the child's father has deceased
PARDEC	Dummy=1 if both the child's parents have deceased
M_PROPMIGF	Interaction of MALE x PROPMIGF
M_PROPMIGM	Interaction of MALE x PROPMIGM
M_MIGF	Interaction of MALE x MIGRANTF
M_MIGM	Interaction of MALE x MIGRANTM
1996	Dummy=1 if the respondent was surveyed in 1996 (base year for 1996-1998)
1997	Dummy=1 if the respondent was surveyed in 1997
1998	Dummy=1 if the respondent was surveyed in 1998

Appendix B
Glossary of household level variables

Variable	Description
HEADPS	Dummy = 1 if the household head has some primary school education
HEADSS	Dummy = 1 if the household head has some secondary school education
HEADTERT	Dummy = 1 if the household head has some tertiary education
MALEHH	Dummy = 1 if the household head is male
AGEHH	Age of household head
AGE2HH	Age of household head squared
F_0-6	Number of female children aged 0-5 in the household
F_7-12	Number of female children aged 6-12 in the household
F_13-18	Number of female children aged 13-18 in the household
M_0-6	Number of male children aged 0-5 in the household
M_7-12	Number of male children aged 6-12 in the household
M_13-18	Number of male children aged 13-18 in the household
MDECHH	Dummy=1 if a child in the household's mother is deceased.
FDECHH	Dummy=1 if a child in the household's father is deceased.
PDECHH	Dummy=1 if both parents of a child in the household are deceased.
PROPMIGF	Proportion of households in the cluster sending at least one female migrant
PROPMIGM	Proportion of households in the cluster sending at least one male migrant
MIGRANTF	Dummy = 1 if the household has sent at least one female migrant
MIGRANTM	Dummy = 1 if the household has sent at least one male migrant
AQEQF	Number of females in the household qualifying for the old-age government pension on the basis of age
AQEQM	Number of males in the household qualifying for the old-age government pension on the basis of age
RECPF	Number of females in the household who receive the old-age government pension
RECPM	Number of males in the household who receive the old-age government pension
PRISCH	Number of primary schools in the cluster
SECSCH	Number of secondary schools in the cluster
ROOMSPP	Number of rooms the household occupies (excluding kitchens and bathrooms) per household member
OWNDWEL	Dummy=1 if the household owns the dwelling they occupy
HOUSE	Dummy=1 if the main dwelling of the household is a house.
COMBO	Dummy=1 if the main dwelling of the household is a combination of structures
ROOM	Dummy=1 if the main dwelling of the household is a room
SHACK	Dummy=1 if the main dwelling of the household is a shack
CEMENTW	Dummy=1 if the walls of the main dwelling are cement
IRONW	Dummy=1 if the walls of the main dwelling are corrugated iron
WOODW	Dummy=1 if the walls of the main dwelling are wood
MIXW	Dummy=1 if the walls of the main dwelling are a mixture of materials
WATTLEW	Dummy=1 if the walls of the main dwelling are made of wattle
MUDW	Dummy=1 if the walls of the main dwelling are mud
OTHERW	Dummy=1 if the walls of the main dwelling are other than cement, iron, wood, mixture, wattle or mud: base is brick
CEMENTR	Dummy=1 if the roof of the main dwelling is cement
TILER	Dummy=1 if the roof of the main dwelling is tiled
THATCHR	Dummy=1 if the roof of the main dwelling is thatched
ASBR	Dummy=1 if the roof of the main dwelling is asbestos
OTHERR	Dummy=1 if the roof of the main dwelling is other than cement, tile, thatch or asbestos: base is corrugated iron
PWATER	Dummy=1 if the household has access to piped water

FLUSH	Dummy=1 if the household has access to a flushing toilet
FLUSH_IN	Dummy=1 if the household has access to a flushing toilet inside the dwelling
FLUSH_ON	Dummy=1 if the household has access to a flushing toilet on site
PIT	Dummy=1 if the household has access to a pit toilet
PIT_ON	Dummy=1 if the household has access to a pit toilet on site
RUBBISH	Dummy=1 if household refuge is removed by the local authority or community members at least once per week
ELECTL	Dummy=1 if the household uses electricity as their main lighting source
GASL	Dummy=1 if the household uses gas as their main lighting source
PARRAFL	Dummy=1 if the household uses paraffin as their main lighting source
OTHERL	Dummy=1 if the household uses a main lighting source other than electricity, gas or paraffin: base=candles
ELECTC	Dummy=1 if the household uses electricity as their main cooking source
GASC	Dummy=1 if the household uses gas as their main cooking source
PARRAFC	Dummy=1 if the household uses paraffin as their main lighting source
COALC	Dummy=1 if the household uses coal as their main lighting source
DUNGC	Dummy=1 if the household uses dung as their main lighting source
OTHERC	Dummy=1 if the household uses a main lighting source other than electricity, gas or paraffin: base=wood
ELECTH	Dummy=1 if the household uses electricity as their main heating source
GASH	Dummy=1 if the household uses gas as their main heating source
PARRAFH	Dummy=1 if the household uses paraffin as their main heating source
COALH	Dummy=1 if the household uses coal as their main heating source
DUNGH	Dummy=1 if the household uses dung as their main heating source
OTHERH	Dummy=1 if the household uses a main heating source other than electricity, gas or paraffin: base=wood
CAPE	Dummy = 1 if the household resides in Cape (base province for 1993)
NATAL	Dummy = 1 if the household resides in Natal
TRANSVAAL	Dummy = 1 if the household resides in Transvaal
ORANGEFS	Dummy = 1 if the household resides in Orange Free State
KZNATAL	Dummy = 1 if the household resides in Kwazulu-Natal
KANGWANE	Dummy = 1 if the household resides in Kangwane
QWAQWA	Dummy = 1 if the household resides in Qwa Qwa
GAZANKUL	Dummy = 1 if the household resides in Gazankulu
LEBOWA	Dummy = 1 if the household resides in Lebowa
KWANDEBE	Dummy = 1 if the household resides in Kwandebele
TRANSKEI	Dummy = 1 if the household resides in Transkei
BOPHUTHA	Dummy = 1 if the household resides in Bophuthatswana
VENDA	Dummy = 1 if the household resides in Venda
CISKEI	Dummy = 1 if the household resides in Ciskei
WCAPE	Dummy = 1 if the household resides in Western Cape (base province for 1996-8)
ECAPE	Dummy = 1 if the household resides in Eastern Cape
NCAPE	Dummy = 1 if the household resides in Northern Cape
FSTATE	Dummy = 1 if the household resides in the Free State
KZNATAL	Dummy = 1 if the household resides in Kwazulu-Natal
NWEST	Dummy = 1 if the household resides in North West Province
GAUTENG	Dummy = 1 if the household resides in Gauteng
MPLANGA	Dummy = 1 if the household resides in Mpumalanga
NPROV	Dummy = 1 if the household resides in Northern Province

Appendix C
Glossary of dependent and constructed variables

Variable	**Description**
PCEXP	Log per capita total expenditure of the household
PCEXPE	Residual for OLS first stage regression of PCEXP on household variables
ATTSCH 7-12	Dummy=1 if the child currently attends school and is aged 7-12
ATTSCH 13-18	Dummy=1 if the child currently attends school and is aged 13-18, omitting those who have completed Standard 10 education.
ATTAINED = 0	Highest attained education is no schooling
ATTAINED = 1	Highest attained education is some (but not complete) primary schooling
ATTAINED = 2	Highest attained education is primary school completion
PROPSCH7-12	Proportion of 7-12 year old children attending school, given at least one child in the household is aged 7-12
PROPSCH13-18	Proportion of 13-18 year old children attending school, given at least one child in the household is aged 13-18
PSCHOOL7-12	Dummy=1 if the child has ever attended primary school and is aged 7-12
PSCHOOL13-18	Dummy=1 if the child has ever attended primary school and is aged 13-18

References

Akresh, R. (2004). *Adjusting Household Structure: School Enrolment Impacts of Child Fostering in Burkina Faso*. IZA Working Paper # 1379.

Alderman, H. (1999). *Safety Nets and Income Transfers in South Africa*. World Bank Africa Region Discussion Paper 193335.

Anh, T. S., Knodel, J., Lam D., & Friedman J. (1998). Family Size and Children's Education in Vietnam. *Demography*, **35**(1), 57 – 70.

Becker, G. S., & Lewis, H. G. (1973). Interaction between Quantity and Quality of Children. *Journal of Political Economy,* **81**(2), S279 – S288.

Behrman, J. R., Pollak, A., & Taubman, P. (1989). Child Endowments and the Quantity and Quality of Children. *Journal of Political Economy,* **97**(2), 389 – 419.

Bertrand, M., Mullainathan, S., & Miller, D. (2003). Public Policy and Extended Families: Evidence from Pensions in South Africa. *World Bank Economic Review,* **17**(3), 27 - 50.

Butcher, K. F., & Case A. (1994). The Effect of Sibling Sex Composition on Women's Education and Earnings. *Quarterly Journal of Economics,* **109**(3), 531 – 563.

Case, A. & Deaton, A. (1999). School Inputs and Educational Outcomes in South Africa. *Quarterly Journal of Economics*, **114**(3), 1047 – 1084.

Case, A. & Deaton, A. (1998). Large Cash Transfers to the Elderly in South Africa. *Economic Journal*, **108**, 1330 – 1363.

Chernichovsky, D. (1985). Socioeconomic and Demographic Aspects of School Enrolment and Attendance in Rural Botswana. *Economic Development and Cultural Change*, **33**, 319 – 332.

Conley, D. (2004). *What is the 'True' Effect of Sibship Size and Birth Order on Education? Instrumental Variable Estimates from Exogenous Variation in Fertility*. Mimeo, New York University.

Duflo, E. (2003). Grandmothers and Granddaughters: Old Age Pension and Intrahousehold Allocation in South Africa. *World Bank Economic Review,* **17**(3), 1 – 25.

Edmonds, E., Mammen, K., & Miller, D. (2003). Rearranging the Family? Income Support and Elderly Living Arrangements in a Low-Income Country. *Journal of Human Resources*, **40**(1), 186 – 207.

Gertler, P. & Glewwe, P. (1992). The Willingness to Pay for Education for Daughters in Contrast to Sons: Evidence from Rural Peru. *The World Bank Economic Review,* **6**(1), 171 – 188.

Glewwe, P. & Jacoby, H. G. (1994). Student Achievement and Schooling Choice in Low-Income Countries. *Journal of Human Resources*, **28**(3), 843 - 864.

Gomes, M. (1984). Family Size and Educational Attainment in Kenya. *Population and Development Review*, **10**, 647 – 660.

Goux, D. & Maurin, E. (2004). *The Effects of Overcrowded Housing on Children's Performance at School*. CEPR Discussion Paper # 3818.

Jacoby, H. G. (1994). Borrowing Constraints and Progress through School: Evidence from Peru. *Review of Economics and Statistics,* **76**(1), 151 – 160.

Kaestner, R. (1996). *Are Brothers Really Better: Sibling Sex Composition and Educational Achievement Revisited*. National Bureau of Economic Research Working Paper 5521.

Knodel, J., Havanon, N., & Sittirai, W. (1990). Family Size and the Education of Children in the Context of Rapid Fertility Decline. *Population and Development Review*, **16**, 31 – 62.

Lipton, M. (1980). Migration from Rural Areas of Poor Countries: the Impact on Rural Productivity and Income Distribution. *World Development*, **8**, 1-24.

Lund, F. (1993). State Social Benefits in South Africa. *International Social Security Review*, **46**(1), 5-25.

Malarani, V. (2004). *Family Size and Educational Attainment in Indonesia: A Cohort Perspective*. California Center for Population Research, University of California, Los Angeles.

Nugent, J. B. (1985). The Old-Age Security Motive for Fertility. *Population and Development Review*, **11**(1), 75 - 97.

Nugent, J. B., & Gillaspy, T. R. (1983). Old Age Pensions and Fertility in Rural Areas of Less Developed Countries: Some Evidence from Mexico. *Economic Development and Cultural Change*, **31**(4), 809 – 829.

Psacharopoulos, G. (1997). Child Labour versus Educational Attainment: Some Evidence from Latin America. *Journal of Population Economics,* **10**(4), 377 – 386.

Psacharopoulos, G., & Arriagada, A. M. (1989). The Determinants of Early Age Human Capital Formation: Evidence from Brazil. *Economic Development and Cultural Change*, **37**, 683 – 708.

Qian, N. (2005). *Quantity-Quality: The Positive Effect of Family Size on School Enrolment in China.* Mimeo, Brown University.

Rivers, D., & Vuong, Q. H. (1988). Limited Information Estimators and Exogeneity Tests for Simultaneous Probit Models. *Journal of Econometrics*, **39**, 347 – 366.

SALDRU (1994). *Project for Living Standards and Development, South Africans Rich and Poor: Baseline Economic Statistics*. School of Economics, University of Cape Town, August.

Shavit, Y., & Pierce, J. L. (1991). Sibship Size and Educational Attainment in Nuclear and Extended Families: Arabs and Jews in Israel. *American Sociological Review*, **56** 321 – 330.

Stark, O., Helmenstein, C., & Prskawetz, A. (1997). A Brain Gain with a Brain Drain. *Economics Letters*, **55**, 227 – 234.

Thomas, D. (1996). Education Across Generations in South Africa. *American Economic Review, Papers and Proceedings,* **86**(2), 330 - 334.

Wilson, F. (1972). *Labour in the South African Gold Mines, 1911-1969*. Cambridge: Cambridge University Press, 1972.

Figure 1
Ratio of Rural-Urban Residents by Age Group, Black South Africans 1996-1998

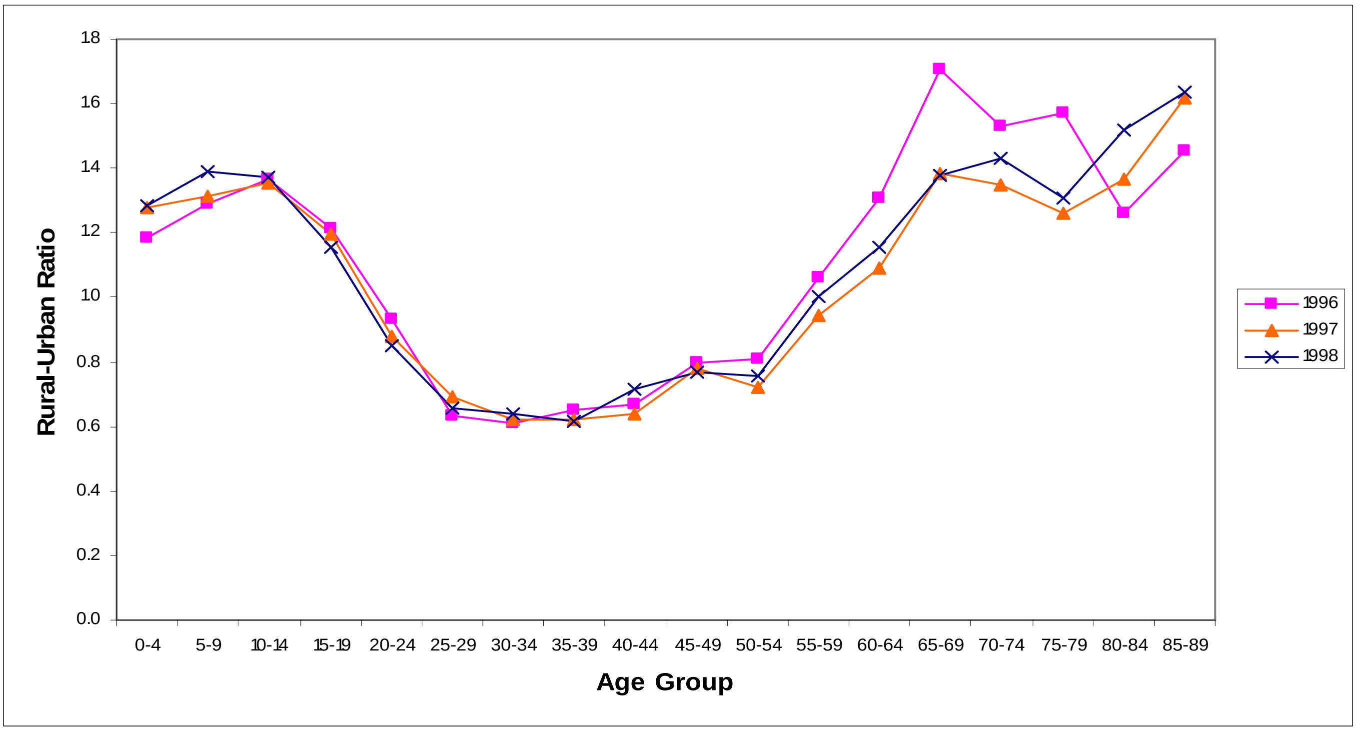

Table 1
Selected descriptive statistics

	SIHS 1993					OHS 1996-8				
	Obs	**Mean**	**Std. Dev.**	**Min**	**Max**	**Obs**	**Mean**	**Std. Dev.**	**Min**	**Max**
Individual level variables										
SIBF_0-6	6053	0.52	0.76	0.00	6.00	46510	1.42	2.07	0.00	15.00
SIBF_7-12	6053	0.76	0.91	0.00	5.00	46510	1.91	2.61	0.00	19.00
SIBF_13-18	6053	0.56	0.74	0.00	4.00	46510	1.61	2.25	0.00	12.00
SIBM_0-6	6053	0.55	0.79	0.00	5.00	46510	1.43	2.08	0.00	14.00
SIBM_7-12	6053	0.77	0.95	0.00	7.00	46510	1.94	2.69	0.00	16.00
SIBM_13-18	6053	0.55	0.77	0.00	5.00	46510	1.52	2.18	0.00	17.00
MOTHERDEC	6053	0.03	0.17	0.00	1.00	46510	0.04	0.19	0.00	1.00
FATHERDEC	6053	0.12	0.33	0.00	1.00	46510	0.15	0.35	0.00	1.00
PARDEC	6053	0.01	0.10	0.00	1.00	46510	0.01	0.12	0.00	1.00
Household level variables										
PROPMIGF	3380	0.13	0.13	0.00	0.63	17683	0.13	0.12	0.00	1.00
PROPMIGM	3380	0.30	0.23	0.00	0.95	17683	0.19	0.14	0.00	0.90
MIGRANTF	3380	0.10	0.30	0.00	1.00	25048	0.16	0.37	0.00	1.00
MIGRANTM	3380	0.17	0.38	0.00	1.00	25048	0.26	0.44	0.00	1.00
AQEQF	3380	0.32	0.55	0.00	6.00	25048	0.86	1.37	0.00	9.00
AQEQM	3380	0.15	0.41	0.00	4.00	25048	0.33	0.69	0.00	5.00
RECPF	3380	0.26	0.45	0.00	3.00	25048	0.71	1.22	0.00	9.00
RECPM	3380	0.12	0.33	0.00	2.00	25048	0.17	0.48	0.00	5.00
Dependent variables										
PCEXP	3380	5.13	0.84	2.28	8.23	25048	4.68	0.98	-1.79	12.43
ATTSCH 7-12 (all years)	3297	0.85	0.35	0.00	1.00	23805	0.93	0.25	0.00	1.00
ATTSCH 7-12 (1996 only)						5754	0.96	0.20	0.00	1.00
ATTSCH 7-12 (1997 only)						11597	0.93	0.26	0.00	1.00
ATTSCH 7-12 (1998 only)						6454	0.92	0.27	0.00	1.00
ATTSCH 13-18 (all years)	2749	0.84	0.36	0.00	1.00	22312	0.90	0.30	0.00	1.00
ATTSCH 13-18 (1996 only)						5252	0.91	0.28	0.00	1.00
ATTSCH 13-18 (1997 only)						10842	0.90	0.30	0.00	1.00
ATTSCH 13-18 (1998 only)						6218	0.89	0.32	0.00	1.00
ATTAINED = 0	80	0.00	0.00	0.00	0.00	773	0.00	0.00	0.00	0.00
ATTAINED = 1	1726	1.00	0.00	1.00	1.00	6005	1.00	0.00	1.00	1.00
ATTAINED = 2	947	2.00	0.00	2.00	2.00	15638	2.00	0.00	2.00	2.00

Table 2
OLS Estimates for PCEXP

	SIHS 1993	OHS 1996-8
HEADPS	0.038***	0.028**
	(0.011)	(0.013)
HEADSS	0.147***	0.142***
	(0.014)	(0.016)
HEADTERT	0.477***	0.520***
	(0.054)	(0.043)
MALEHH	0.065***	-0.006
	(0.024)	(0.013)
AGEHH	-0.018***	-0.008***
	(0.005)	(0.002)
AGE2HH	0.000***	0.000***
	(0.000)	(0.000)
ROOMSPP	0.331***	0.229***
	(0.047)	(0.027)
DWELPP		0.571***
		(0.031)
OWNDWEL	0.065**	0.130***
	(0.031)	(0.023)
HOUSE	-0.046	-0.023
	(0.033)	(0.018)
COMBO	-0.069**	
	(0.034)	
ROOM		-0.038
		(0.041)
SHACK	0.069	-0.039
	(0.066)	(0.036)
CEMENTW	0.031	-0.018
	(0.035)	(0.019)
IRONW	-0.281***	0.018
	(0.079)	(0.037)
WOODW	0.085	-0.034
	(0.101)	(0.059)
MIXW	-0.041	-0.092***
	(0.039)	(0.022)
WATTLEW	0.058	-0.098
	(0.050)	(0.062)
MUDW	-0.136***	-0.114***
	(0.041)	(0.019)
OTHERW	-0.013	-0.082
	(0.111)	(0.064)
CEMENTR	-0.063	-0.038
	(0.089)	(0.045)
TILER	0.154	0.332***
	(0.104)	(0.044)
THATCHR	0.019	-0.076***
	(0.039)	(0.019)
ASBR	0.117***	0.022
	(0.047)	(0.038)
OTHERR	-0.060	0.024
	(0.068)	(0.057)
PWATER	0.015	0.069***
	(0.024)	(0.015)
FLUSH	0.693***	0.179*
	(0.171)	(0.108)
FLUSH_IN	-0.226	0.007
	(0.181)	(0.108)
FLUSH_ON	-0.238	-0.061

	(0.185)	(0.110)
PIT	0.256***	0.039
	(0.096)	(0.026)
PIT_ON	-0.079	0.046*
	(0.095)	(0.024)
RUBBISH		0.240***
		(0.033)
ELECTL	0.081**	0.056***
	(0.040)	(0.017)
GASL	0.380**	0.038
	(0.190)	(0.079)
PARRAFL	0.011	0.069***
	(0.027)	(0.016)
OTHERL	0.133	-0.002
	(0.105)	(0.078)
ELECTC	0.532***	0.296***
	(0.073)	(0.047)
GASC	0.677***	0.429***
	(0.074)	(0.063)
PARRAFC	0.248***	0.138***
	(0.039)	(0.034)
COALC	0.325***	0.117**
	(0.078)	(0.056)
DUNGC	-0.211**	0.103
	(0.094)	(0.140)
OTHERC	-0.041	0.494***
	(0.169)	(0.176)
ELECTH	0.074	0.089*
	(0.074)	(0.047)
GASH	-0.133	0.084
	(0.120)	(0.074)
PARRAFH	-0.022	0.057*
	(0.050)	(0.034)
COALH	-0.103	0.011
	(0.075)	(0.053)
DUNGH	0.051	-0.155
	(0.090)	(0.147)
OTHERH	-0.114***	0.002
	(0.043)	(0.066)
1997		-0.258***
		(0.018)
1998		-0.277***
		(0.025)
Constant	4.761***	3.970***
	(0.139)	(0.062)
Observations	18946	129580
R^2	0.42	0.28

Robust standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%

Table 3
Probit Marginal Effects for ATTSCH

	SIHS 1993		**OHS 1996-98**	
	(1) **Marginal Effects 7-12yo**	**(2)** **Marginal Effects 13-18yo**	**(3)** **Marginal Effects 7-12yo**	**(4)** **Marginal Effects 13-18yo**
MALE	-0.017	0.032	-0.006	0.008
	(0.022)	(0.026)	(0.004)	(0.006)
AGE	0.261***	0.039	0.115***	0.076***
	(0.044)	(0.086)	(0.011)	(0.023)
AGESQ	-0.012***	-0.002	-0.005***	-0.003***
	(0.002)	(0.003)	(0.001)	(0.001)
SIBF_0-6	0.013	-0.006	-0.001	-0.007***
	(0.012)	(0.013)	(0.002)	(0.002)
M_SIBF_0-6	-0.017	-0.004	-0.002	0.007**
	(0.014)	(0.017)	(0.002)	(0.003)
SIBF_7-12	-0.012	0.028**	0.002	0.004*
	(0.012)	(0.013)	(0.002)	(0.002)
M_SIBF_7-12	0.014	-0.022	0.001	0.001
	(0.015)	(0.017)	(0.002)	(0.003)
SIBF_13-18	0.015	-0.008	-0.000	0.006***
	(0.013)	(0.016)	(0.002)	(0.002)
M_SIBF_13-18	-0.002	0.002	0.002	-0.004
	(0.017)	(0.021)	(0.002)	(0.003)
SIBM_0-6	-0.004	-0.017	0.001	-0.005***
	(0.010)	(0.012)	(0.002)	(0.002)
M_SIBM_0-6	-0.002	-0.010	-0.002	0.002
	(0.012)	(0.016)	(0.002)	(0.003)
SIBM_7-12	-0.016	0.006	-0.001	0.006***
	(0.012)	(0.013)	(0.002)	(0.002)
M_SIBM_7-12	0.000	0.005	0.003	-0.005
	(0.014)	(0.015)	(0.002)	(0.003)
SIBM_13-18	0.004	-0.016	0.001	0.002
	(0.012)	(0.014)	(0.002)	(0.002)
M_SIBM_13-18	0.013	0.009	-0.001	0.002
	(0.017)	(0.019)	(0.002)	(0.003)
MOTHERDEC	-0.063	-0.073	0.025***	-0.011
	(0.072)	(0.056)	(0.009)	(0.013)
FATHERDEC	-0.010	-0.036	0.013**	-0.006
	(0.028)	(0.025)	(0.005)	(0.006)
PARDEC	0.079*	-0.038	-0.115**	-0.027
	(0.044)	(0.077)	(0.048)	(0.023)
MIGRANTF	0.002	-0.036	0.010*	0.012*
	(0.027)	(0.034)	(0.006)	(0.007)
M_MIGF	0.016	-0.019	-0.007	-0.016
	(0.032)	(0.043)	(0.009)	(0.011)
MIGRANTM	-0.032	-0.004	0.015***	0.007
	(0.025)	(0.025)	(0.005)	(0.006)
M_MIGM	0.031	0.023	-0.004	0.014*
	(0.025)	(0.032)	(0.007)	(0.008)
AGEQF	0.022	0.038*	-0.002	0.008**
	(0.018)	(0.022)	(0.003)	(0.003)
M_AGEQF	-0.036	-0.058**	0.002	-0.006
	(0.022)	(0.027)	(0.003)	(0.004)
AGEQM	0.005	-0.067***	0.003	-0.008
	(0.022)	(0.023)	(0.004)	(0.005)
M_AGEQM	-0.020	-0.002	-0.010*	0.002
	(0.027)	(0.034)	(0.005)	(0.007)
PCEXP	0.084***	0.090***	0.037***	0.046***
	(0.023)	(0.026)	(0.005)	(0.006)
PRISCH	-0.000	0.001		
	(0.005)	(0.005)		
SECSCH		-0.009		

		(0.013)		
Y97			-0.020***	0.003
			(0.005)	(0.006)
Y98			-0.036***	-0.053***
			(0.012)	(0.013)
PCEXPE	-0.080***	-0.089***	-0.037***	-0.040***
	(0.026)	(0.029)	(0.006)	(0.007)
Observations	3191	2651	22373	20944
Pseudo R^2	0.13	0.11	0.08	0.08
Log Pseudo-Likelihood	-1138.07	-1019.69	-5092.42	-6250.86
Wald χ^2	269.81	207.75	729.28	980.89
Degrees of Freedom	41.00	41.00	38.00	38.00

Robust standard errors in parentheses

* significant at 10%; ** significant at 5%; *** significant at 1%

Regressions also control for a set of pre-1994 province dummies for SIHS1993 (reference category household resides is Cape Province) and a set of post-1994 province dummies for OHS 1996-98 (reference category household resides in Western Cape)

PCEXPE denote predicted value of PCEXP from first stage regression (see Table 2)

Table 4
Probit Marginal Effects for ATTSCH
Using RECPF and RECPM

	SIHS 1993		OHS 1996-98	
	(1) **Marginal Effects 7-12 year olds**	**(2)** **Marginal Effects 13-18 year olds**	**(3)** **Marginal Effects 7-12 year olds**	**(4)** **Marginal Effects 13-18 year olds**
MALE	-0.022	0.032	-0.006	0.007
	(0.022)	(0.027)	(0.004)	(0.006)
AGE	0.261***	0.051	0.115***	0.076***
	(0.045)	(0.086)	(0.011)	(0.023)
AGESQ	-0.012***	-0.003	-0.005***	-0.003***
	(0.002)	(0.003)	(0.001)	(0.001)
SIBF_0-6	0.012	-0.007	-0.001	-0.007***
	(0.013)	(0.013)	(0.002)	(0.002)
M_SIBF_0-6	-0.015	-0.005	-0.002	0.008**
	(0.014)	(0.017)	(0.002)	(0.003)
SIBF_7-12	-0.012	0.030**	0.002	0.004
	(0.012)	(0.012)	(0.002)	(0.002)
M_SIBF_7-12	0.013	-0.024	0.000	0.001
	(0.015)	(0.017)	(0.002)	(0.003)
SIBF_13-18	0.015	-0.008	-0.000	0.006***
	(0.013)	(0.016)	(0.002)	(0.002)
M_SIBF_13-18	-0.002	-0.002	0.002	-0.004
	(0.017)	(0.021)	(0.002)	(0.003)
SIBM_0-6	-0.003	-0.016	0.001	-0.005***
	(0.010)	(0.012)	(0.002)	(0.002)
M_SIBM_0-6	-0.003	-0.014	-0.002	0.002
	(0.012)	(0.017)	(0.002)	(0.003)
SIBM_7-12	-0.018	0.005	-0.001	0.006***
	(0.012)	(0.013)	(0.002)	(0.002)
M_SIBM_7-12	-0.000	0.005	0.003	-0.005
	(0.015)	(0.015)	(0.002)	(0.003)
SIBM_13-18	0.005	-0.015	0.001	0.002
	(0.012)	(0.014)	(0.002)	(0.002)
M_SIBM_13-18	0.012	0.009	-0.001	0.002
	(0.017)	(0.018)	(0.002)	(0.003)
MOTHERDEC	-0.068	-0.070	0.025***	-0.011
	(0.072)	(0.056)	(0.009)	(0.013)
FATHERDEC	-0.008	-0.032	0.013**	-0.006
	(0.027)	(0.025)	(0.005)	(0.006)
PARDEC	0.080*	-0.040	-0.114**	-0.027
	(0.042)	(0.078)	(0.048)	(0.023)
MIGRANTF	-0.007	-0.038	0.010*	0.012*
	(0.029)	(0.034)	(0.006)	(0.007)
M_MIGF	0.023	-0.021	-0.007	-0.016
	(0.030)	(0.044)	(0.009)	(0.011)
MIGRANTM	-0.029	-0.009	0.015***	0.006
	(0.025)	(0.026)	(0.005)	(0.006)
M_MIGM	0.024	0.026	-0.003	0.014*
	(0.026)	(0.031)	(0.007)	(0.008)
RECPF	0.057***	0.046*	0.000	0.010***
	(0.021)	(0.024)	(0.003)	(0.004)
M_RECPF	-0.060**	-0.036	0.000	-0.007
	(0.026)	(0.032)	(0.003)	(0.004)
RECPM	-0.016	-0.040	-0.001	-0.007
	(0.024)	(0.034)	(0.005)	(0.006)
M_RECPM	0.035	-0.030	-0.007	0.001
	(0.030)	(0.046)	(0.006)	(0.007)
PRISCH	-0.000	0.002		
	(0.005)	(0.006)		
SECSCH		-0.011		
		(0.013)		

PCEXP	0.088***	0.092***	0.037***	0.046***
	(0.023)	(0.027)	(0.005)	(0.006)
Y97			-0.020***	0.003
			(0.005)	(0.006)
Y98			-0.039***	-0.055***
			(0.011)	(0.014)
Residuals	-0.083***	-0.091***	-0.037***	-0.040***
	(0.026)	(0.029)	(0.006)	(0.007)
Observations	2901	2422	22373	20944
Pseudo R^2	0.13	0.11	0.08	0.08
Log Pseudo-Likelihood	-1033.38	-942.61	-5093.92	-6249.97
Wald χ^2	239.59	195.44	724.68	980.77
Degrees of Freedom	42.00	43.00	38.00	38.00

Robust standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%
Regressions also control for a set of pre-1994 province dummies for SIHS1993 (reference category household resides is Cape Province) and a set of post-1994 province dummies for OHS 1996-98 (reference category household resides in Western Cape)
Residuals denote predicted value of PCEXP from first stage regression (see Table 2)

Table 5
Probit Marginal Effects for ATTSCH
Proportion of Households in the Cluster Sending Migrants

	SIHS 1993		OHS 1996-97	
	(1) Marginal Effects 7-12yo	(2) Marginal Effects 13-18yo	(3) Marginal Effects 7-12yo	(4) Marginal Effects 13-18yo
MALE	-0.029	0.029	-0.010	0.007
	(0.027)	(0.031)	(0.008)	(0.011)
AGE	0.265***	0.040	0.123***	0.069**
	(0.045)	(0.086)	(0.013)	(0.027)
AGESQ	-0.012***	-0.002	-0.006***	-0.003***
	(0.002)	(0.003)	(0.001)	(0.001)
SIBF_0-6	0.012	-0.007	0.000	-0.020***
	(0.012)	(0.013)	(0.003)	(0.004)
M_SIBF_0-6	-0.017	-0.002	-0.005	0.025***
	(0.014)	(0.017)	(0.004)	(0.005)
SIBF_7-12	-0.012	0.028**	0.005	0.012***
	(0.012)	(0.013)	(0.004)	(0.004)
M_SIBF_7-12	0.014	-0.023	-0.004	-0.004
	(0.015)	(0.017)	(0.005)	(0.006)
SIBF_13-18	0.013	-0.011	0.000	0.017***
	(0.013)	(0.016)	(0.004)	(0.005)
M_SIBF_13-18	-0.001	0.004	0.005	-0.015**
	(0.017)	(0.021)	(0.004)	(0.007)
SIBM_0-6	-0.003	-0.017	-0.003	-0.015***
	(0.010)	(0.012)	(0.004)	(0.004)
M_SIBM_0-6	-0.002	-0.010	0.003	0.012**
	(0.012)	(0.016)	(0.004)	(0.005)
SIBM_7-12	-0.016	0.006	-0.003	0.019***
	(0.012)	(0.013)	(0.004)	(0.004)
M_SIBM_7-12	0.000	0.006	0.001	-0.019***
	(0.014)	(0.015)	(0.005)	(0.006)
SIBM_13-18	0.003	-0.017	0.002	0.003
	(0.012)	(0.014)	(0.004)	(0.005)
M_SIBM_13-18	0.012	0.012	-0.003	0.008
	(0.017)	(0.019)	(0.004)	(0.007)
MOTHERDEC	-0.073	-0.072	0.007	-0.010
	(0.073)	(0.056)	(0.013)	(0.016)
FATHERDEC	-0.009	-0.037	0.000	-0.008
	(0.028)	(0.025)	(0.006)	(0.007)
PARDEC	0.080*	-0.036	-0.048	-0.027
	(0.042)	(0.076)	(0.036)	(0.027)
PROPMIGF	0.081	0.077	0.054**	0.049
	(0.084)	(0.099)	(0.025)	(0.030)
M_PROPMIGF	0.120	-0.010	-0.025	-0.045
	(0.114)	(0.127)	(0.030)	(0.037)
PROPMIGM	-0.114**	-0.075	-0.019	-0.042
	(0.058)	(0.064)	(0.021)	(0.026)
M_PROPMIGM	0.006	0.008	0.034	0.030
	(0.067)	(0.077)	(0.026)	(0.032)
AGEQF	0.021	0.037*	-0.002	0.006
	(0.018)	(0.021)	(0.005)	(0.008)
M_AGEQF	-0.037*	-0.063**	0.003	-0.009
	(0.022)	(0.027)	(0.007)	(0.010)
AGEQM	0.001	-0.067***	-0.012	-0.011
	(0.021)	(0.023)	(0.008)	(0.010)
M_AGEQM	-0.016	0.006	0.010	-0.001
	(0.027)	(0.034)	(0.010)	(0.014)
PCEXP	0.073***	0.084***	0.032***	0.035***
	(0.024)	(0.027)	(0.006)	(0.007)
PRISCH	-0.001	0.000		

	(0.005)	(0.006)		
SECSCH		-0.007		
		(0.014)		
Y97			-0.012**	0.010
			(0.005)	(0.007)
PCEXPE	-0.069***	-0.083***	-0.037***	-0.032***
	(0.027)	(0.030)	(0.007)	(0.007)
Observations	2901	2422	16139	14960
Pseudo R^2	0.13	0.11	0.09	0.09
Log Pseudo-Likelihood	-1032.09	-938.86	-3449.71	-4267.47
Wald χ^2	244.70	195.26	551.30	727.33
Degrees of Freedom	42.00	43.00	37.00	37.00

Robust standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%
Regressions also control for a set of pre-1994 province dummies for SIHS1993 (reference category household resides is Cape Province) and a set of post-1994 province dummies for OHS 1996-98 (reference category household resides in Western Cape)
PCEXPE denote predicted value of PCEXP from first stage regression (see Table 2)

Table 6: Ordered Probit Estimates for ATTAINED
Children Aged 13-18

	SIHS 1993				OHS 1996-98			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Coefficients	Marginal Effects			Coefficients	Marginal Effects		
		Attained=0	Attained=1	Attained=2		Attained=0	Attained=1	Attained=2
MALE	-0.358***	0.021	0.076	-0.097	-0.332***	0.023	0.078	-0.101
	(0.104)				(0.029)			
AGE	0.186	-0.011	-0.040	0.051	1.298***	-0.089	-0.305	0.394
	(0.353)				(0.114)			
AGESQ	0.003	-0.000	-0.001	0.001	-0.034***	0.002	0.008	-0.010
	(0.011)				(0.004)			
SIBF_0-6	0.054	-0.003	-0.012	0.015	-0.022**	0.001	0.005	-0.007
	(0.050)				(0.010)			
M_SIBF_0-6	0.024	-0.001	-0.005	0.006	0.012	-0.001	-0.003	0.003
	(0.066)				(0.013)			
SIBF_7-12	0.127***	-0.007	-0.027	0.035	0.007	-0.000	-0.002	0.002
	(0.049)				(0.014)			
M_SIBF_7-12	-0.102	0.006	0.022	-0.028	-0.002	0.000	0.001	-0.001
	(0.063)				(0.016)			
SIBF_13-18	0.087	-0.005	-0.019	0.024	-0.019	0.001	0.004	-0.006
	(0.060)				(0.013)			
M_SIBF_13-18	-0.097	0.006	0.021	-0.026	0.020	-0.001	-0.005	0.006
	(0.080)				(0.015)			
SIBM_0-6	-0.014	0.001	0.003	-0.004	-0.003	0.000	0.001	-0.001
	(0.051)				(0.011)			
M_SIBM_0-6	-0.038	0.002	0.008	-0.010	-0.004	0.000	0.001	-0.001
	(0.070)				(0.013)			
SIBM_7-12	0.035	-0.002	-0.008	0.010	0.017	-0.001	-0.004	0.005
	(0.051)				(0.011)			
M_SIBM_7-12	0.054	-0.003	-0.012	0.015	-0.011	0.001	0.003	-0.003
	(0.063)				(0.013)			
SIBM_13-18	0.040	-0.002	-0.009	0.011	0.020	-0.001	-0.005	0.006
	(0.055)				(0.013)			
M_SIBM_13-18	0.043	-0.002	-0.009	0.012	-0.017	0.001	0.004	-0.005
	(0.071)				(0.015)			
MOTHERDEC	-0.112	0.007	0.023	-0.030	-0.056	0.004	0.013	-0.017
	(0.156)				(0.060)			
FATHERDEC	0.100	-0.006	-0.022	0.028	-0.033	0.002	0.008	-0.010
	(0.089)				(0.027)			
PARDEC	-0.460*	0.037	0.078	-0.115	-0.130	0.010	0.031	-0.040

	(0.263)				(0.101)			
MIGRANTF	-0.105	0.006	0.022	-0.028	0.022	-0.002	-0.005	0.007
	(0.121)				(0.038)			
M_MIGF	0.396***	-0.018	-0.095	0.113	-0.036	0.003	0.009	-0.011
	(0.152)				(0.048)			
MIGRANTM	-0.060	0.004	0.013	-0.016	0.032	-0.002	-0.008	0.010
	(0.101)				(0.032)			
M_MIGM	-0.007	0.000	0.001	-0.002	-0.020	0.001	0.005	-0.006
	(0.141)				(0.040)			
AGEQF	0.039	-0.002	-0.008	0.011	0.014	-0.001	-0.003	0.004
	(0.079)				(0.017)			
M_AGEQF	-0.122	0.007	0.026	-0.033	0.014	-0.001	-0.003	0.004
	(0.104)				(0.021)			
AGEQM	-0.085	0.005	0.018	-0.023	-0.008	0.001	0.002	-0.002
	(0.117)				(0.028)			
M_AGEQM	0.043	-0.002	-0.009	0.012	-0.009	0.001	0.002	-0.003
	(0.161)				(0.033)			
PRISCH	0.023	-0.001	-0.005	0.006				
	(0.025)							
SECSCH	-0.105**	0.006	0.023	-0.029				
	(0.053)							
PCEXP	1.579***	-0.092	-0.339	0.431	0.522***	-0.036	-0.122	0.158
	(0.111)				(0.031)			
Y97					0.071**	-0.005	-0.017	0.021
					(0.028)			
Y98					0.050	-0.003	-0.012	0.015
					(0.059)			
PCEXPE	-1.461***	0.085	0.314	-0.399	-0.459***	0.031	0.108	-0.139
	(0.121)	(1.097)	(2.248)	(1.151)	(0.034)			
τ_1	9.647***				12.082***			
	(2.758)				(0.889)			
τ_2	12.564***				13.542***			
	(2.775)				(0.891)			
Observations	2425	2425	2425	2425	21042	21042	21042	21042
Pseudo R^2	0.24				0.11			
Log Pseudo-Likelihood	-1397.17				-13412.08			
Wald χ^2	462.71				2612.87			
Degrees of Freedom	43.00				38.00			

Robust standard errors in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. Regressions also control for a set of pre-1994 province dummies for SIHS1993 (reference category household resides is Cape Province) and a set of post-1994 province dummies for OHS 1996-98 (reference category household resides in Western Cape). PCEXPE denote predicted value of PCEXP from first stage regression (see Table 2)

Table 7: Ordered Probit Estimates for ATTAINED Children Aged 13-18
Using RECPF and RECPM

	SIHS 1993				OHS 1996-98			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Coefficients	Marginal Effects			Coefficients	Marginal Effects		
		Attained=0	Attained=1	Attained=2		Attained=0	Attained=1	Attained=2
MALE	-0.353***	0.021	0.075	-0.096	-0.331***	0.023	0.077	-0.100
	(0.105)				(0.029)			
AGE	0.233	-0.014	-0.050	0.064	1.299***	-0.089	-0.305	0.394
	(0.351)				(0.114)			
AGESQ	0.002	-0.000	-0.000	0.001	-0.034***	0.002	0.008	-0.010
	(0.011)				(0.004)			
SIBF_0-6	0.063	-0.004	-0.014	0.017	-0.023**	0.002	0.005	-0.007
	(0.050)				(0.010)			
M_SIBF_0-6	0.012	-0.001	-0.002	0.003	0.013	-0.001	-0.003	0.004
	(0.067)				(0.013)			
SIBF_7-12	0.126***	-0.007	-0.027	0.034	0.007	-0.000	-0.002	0.002
	(0.049)				(0.013)			
M_SIBF_7-12	-0.097	0.006	0.021	-0.026	-0.002	0.000	0.000	-0.001
	(0.063)				(0.016)			
SIBF_13-18	0.089	-0.005	-0.019	0.024	-0.020	0.001	0.005	-0.006
	(0.060)				(0.013)			
M_SIBF_13-18	-0.108	0.006	0.023	-0.029	0.023	-0.002	-0.005	0.007
	(0.080)				(0.015)			
SIBM_0-6	-0.010	0.001	0.002	-0.003	-0.003	0.000	0.001	-0.001
	(0.051)				(0.011)			
M_SIBM_0-6	-0.040	0.002	0.009	-0.011	-0.004	0.000	0.001	-0.001
	(0.070)				(0.013)			
SIBM_7-12	0.027	-0.002	-0.006	0.007	0.016	-0.001	-0.004	0.005
	(0.051)				(0.011)			
M_SIBM_7-12	0.050	-0.003	-0.011	0.014	-0.010	0.001	0.002	-0.003
	(0.064)				(0.013)			
SIBM_13-18	0.036	-0.002	-0.008	0.010	0.018	-0.001	-0.004	0.005
	(0.055)				(0.013)			
M_SIBM_13-18	0.047	-0.003	-0.010	0.013	-0.015	0.001	0.003	-0.004
	(0.071)				(0.015)			
MOTHERDEC	-0.122	0.008	0.025	-0.033	-0.056	0.004	0.013	-0.017
	(0.153)				(0.060)			
FATHERDEC	0.116	-0.006	-0.026	0.032	-0.033	0.002	0.008	-0.010
	(0.088)				(0.027)			
PARDEC	-0.446*	0.035	0.076	-0.112	-0.128	0.010	0.030	-0.040

	(0.263)				(0.101)			
MIGRANTF	-0.125	0.008	0.026	-0.034	0.022	-0.002	-0.005	0.007
	(0.119)				(0.038)			
M_MIGF	0.386**	-0.018	-0.092	0.110	-0.036	0.002	0.008	-0.011
	(0.153)				(0.048)			
MIGRANTM	-0.070	0.004	0.015	-0.019	0.033	-0.002	-0.008	0.010
	(0.102)				(0.032)			
M_MIGM	0.007	-0.000	-0.001	0.002	-0.020	0.001	0.005	-0.006
	(0.142)				(0.040)			
RECPF	0.126	-0.007	-0.027	0.034	0.034*	-0.002	-0.008	0.010
	(0.092)				(0.019)			
M_RECPF	-0.027	0.002	0.006	-0.007	-0.006	0.000	0.001	-0.002
	(0.121)				(0.022)			
RECPM	0.145	-0.008	-0.032	0.040	-0.011	0.001	0.003	-0.003
	(0.110)				(0.030)			
M_RECPM	-0.171	0.011	0.034	-0.045	0.006	-0.000	-0.001	0.002
	(0.157)				(0.036)			
PCEXP	1.606***	-0.094	-0.344	0.438	0.521***	-0.036	-0.122	0.158
	(0.112)				(0.031)			
PRISCH	0.028	-0.002	-0.006	0.008				
	(0.025)							
SECSCH	-0.116**	0.007	0.025	-0.032				
	(0.053)							
Y97					0.066**	-0.004	-0.015	0.020
					(0.028)			
Y98					0.038	-0.003	-0.009	0.011
					(0.059)			
PCEXPE	-1.489***	0.087	0.319	-0.406	-0.460***	0.031	0.108	-0.139
	(0.122)				(0.034)			
τ_1	10.203***				12.081***			
	(2.743)				(0.889)			
τ_2	13.124***				13.542***			
	(2.761)				(0.892)			
Observations	2425	2425	2425	2425	21042	21042	21042	21042
Pseudo R^2	0.24				0.11			
Log Pseudo-Likelihood	-1395.66				-13410.44			
Wald χ^2	472.33				2620.34			
Degrees of Freedom	43.00				38.00			

Robust standard errors in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. Regressions also control for a set of pre-1994 province dummies for SIHS1993 (reference category household resides is Cape Province) and a set of post-1994 province dummies for OHS 1996-98 (reference category household resides in Western Cape). Residuals denote predicted value of PCEXP from first stage regression (see Table 2).

Table 8: Ordered Probit Estimates for ATTAINED Children Aged 13-18
Proportion of Households in the Cluster Sending Migrants

	SIHS1993				OHS 1996-97			
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
	Coefficients		**Marginal Effects**		**Coefficients**		**Marginal Effects**	
		Attained=0	**Attained=1**	**Attained=2**		**Attained=0**	**Attained=1**	**Attained=2**
MALE	-0.422***	0.025	0.089	0.089	-0.275***	0.019	0.064	-0.083
	(0.128)				(0.058)			
AGE	0.273	-0.016	-0.058	-0.058	1.221***	-0.084	-0.284	0.369
	(0.353)				(0.137)			
AGESQ	0.001	-0.000	-0.000	-0.000	-0.032***	0.002	0.007	-0.010
	(0.011)				(0.004)			
SIBF_0-6	0.052	-0.003	-0.011	-0.011	-0.067***	0.005	0.016	-0.020
	(0.050)				(0.022)			
M_SIBF_0-6	0.017	-0.001	-0.004	-0.004	0.052*	-0.004	-0.012	0.016
	(0.066)				(0.027)			
SIBF_7-12	0.124**	-0.007	-0.027	-0.027	0.053**	-0.004	-0.012	0.016
	(0.049)				(0.023)			
M_SIBF_7-12	-0.102	0.006	0.022	0.022	-0.039	0.003	0.009	-0.012
	(0.063)				(0.030)			
SIBF_13-18	0.070	-0.004	-0.015	-0.015	-0.020	0.001	0.005	-0.006
	(0.059)				(0.030)			
M_SIBF_13-18	-0.072	0.004	0.015	0.015	0.030	-0.002	-0.007	0.009
	(0.080)				(0.034)			
SIBM_0-6	-0.011	0.001	0.002	0.002	-0.016	0.001	0.004	-0.005
	(0.050)				(0.023)			
M_SIBM_0-6	-0.034	0.002	0.007	0.007	0.006	-0.000	-0.001	0.002
	(0.070)				(0.028)			
SIBM_7-12	0.032	-0.002	-0.007	-0.007	0.046**	-0.003	-0.011	0.014
	(0.051)				(0.023)			
M_SIBM_7-12	0.042	-0.002	-0.009	-0.009	-0.040	0.003	0.009	-0.012
	(0.062)				(0.031)			
SIBM_13-18	0.045	-0.003	-0.010	-0.010	0.038	-0.003	-0.009	0.012
	(0.055)				(0.028)			
M_SIBM_13-18	0.039	-0.002	-0.008	-0.008	-0.054	0.004	0.013	-0.016
	(0.070)				(0.036)			
MOTHERDEC	-0.120	0.008	0.025	0.025	-0.065	0.005	0.015	-0.020
	(0.155)				(0.074)			
FATHERDEC	0.109	-0.006	-0.024	-0.024	-0.058*	0.004	0.014	-0.018
	(0.089)				(0.033)			
PARDEC	-0.460*	0.037	0.078	0.078	-0.140	0.011	0.033	-0.043

	(0.263)				(0.120)			
PROPMIGF	1.166***	-0.068	-0.249	-0.249	0.281*	-0.019	-0.066	0.085
	(0.444)				(0.162)			
M_PROPMIGF	-0.099	0.006	0.021	0.021	-0.188	0.013	0.044	-0.057
	(0.519)				(0.181)			
PROPMIGM	-0.277	0.016	0.059	0.059	-0.125	0.009	0.029	-0.038
	(0.291)				(0.145)			
M_PROPMIGM	0.410	-0.024	-0.088	-0.088	-0.202	0.014	0.047	-0.061
	(0.338)				(0.168)			
AGEQF	0.027	-0.002	-0.006	-0.006	-0.012	0.001	0.003	-0.004
	(0.078)				(0.036)			
M_AGEQF	-0.085	0.005	0.018	0.018	0.035	-0.002	-0.008	0.011
	(0.104)				(0.047)			
AGEQM	-0.094	0.005	0.020	0.020	-0.144**	0.010	0.033	-0.043
	(0.110)				(0.061)			
M_AGEQM	0.001	-0.000	-0.000	-0.000	0.103	-0.007	-0.024	0.031
	(0.156)				(0.072)			
PRISCH	0.025	-0.001	-0.005	-0.005				
	(0.025)							
SECSCH	-0.072	0.004	0.015	0.015				
	(0.055)							
PCEXP	1.611***	-0.094	-0.344	-0.344	0.459***	-0.032	-0.107	0.138
	(0.114)				(0.036)			
Y97					0.112***	-0.008	-0.026	0.034
					(0.035)			
PCEXPE	-1.517***	0.088	0.324	0.324	-0.410***	0.028	0.096	-0.124
	(0.124)				(0.039)			
τ_1	10.453***				11.084***			
	(2.760)				(1.070)			
τ_2	13.378***				12.526***			
	(2.777)				(1.074)			
Observations	2425	2425	2425	2425	15025	15025	15025	15025
Pseudo R^2	0.24				0.11			
Log Pseudo-Likelihood	-1391.43				-9549.77			
Wald χ^2	471.02				1905.48			
Degrees of Freedom	43.00				37.00			

Robust standard errors in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.Regressions also control for a set of pre-1994 province dummies for SIHS1993 (reference category household resides is Cape Province) and a set of post-1994 province dummies for OHS 1996-98 (reference category household resides in Western Cape). PCEXPE denote predicted value of PCEXP from first stage regression (see Table 2).