Home background and schooling outcomes in South Africa: Insights from the National Income Dynamics Study

HELEEN HOFMEYR

Stellenbosch Economic Working Papers: WP01/2018

www.ekon.sun.ac.za/wpapers/2018/wp012018

January 2018

KEYWORDS: home background, family structure, educational outcomes
JEL: I24, J12, J13

ReSEP (Research on Socio-Economic Policy)
http://resep.sun.ac.za

DEPARTMENT OF ECONOMICS
UNIVERSITY OF STELLENBOSCH

SOUTH AFRICA

A WORKING PAPER OF THE DEPARTMENT OF ECONOMICS AND THE
BUREAU FOR ECONOMIC RESEARCH AT THE UNIVERSITY OF STELLENBOSCH

www.ekon.sun.ac.za/wpapers

Home background and schooling outcomes in South Africa: Insights from the National Income Dynamics Study

Heleen Hofmeyr[1]

Abstract

Patterns of poverty and inequality in South Africa are largely sustained by differential educational outcomes of children across different strata of society. Most of these differences in educational outcomes are attributed to large differences in the quality of education received by children. It is the intention of this paper to add to our understanding of the determinants of educational outcomes in South Africa by investigating the role of the family in determining these heterogeneous educational outcomes. This is done by estimating the relationship between home background and schooling outcomes for a sample of South African youths. The analysis in this paper suggests a strong correlation between home background and the educational outcomes of the sample at hand. Broadly, the co-residence of biological parents in the household and the education attained by parents are found to be positively associated with educational outcomes of children.

Keywords: home background, family structure, educational outcomes
JEL codes: I24, J12, J13

[1] RESEP, Department of Economics, Stellenbosch University. Private Bag X1, 7602, Matieland, South Africa. Email: heleenhofmeyr@sun.ac.za.

1. Introduction

People differ widely in their beliefs about the extent to which family influences one's outcomes in life. While few would contest that family background plays an important role in determining one's life chances, there is little agreement about the exact mechanisms through which family impacts on children's outcomes, as well as the relative sizes of these different effects. Such questions are especially important South Africa. The apartheid regime impacted heavily on all facets of society, including the family. This fact alone makes the family a likely institution through which apartheid injustices continue to plague South African society. Furthermore, the country remains one of the most unequal in the world. Questions about what determines one's life chances are all the more pressing in a society with such vast differences between the life chances of its members.

Differential educational outcomes of children across different strata of society has been identified as one of the most important factors that perpetuate patterns of poverty and inequality in South Africa. Most of these differences are attributed to differences across social classes in the quality of education received by children. It is the intention of this paper to investigate the role of the family in determining these heterogeneous educational outcomes. This is done by estimating the relationship between home background and schooling outcomes for a sample of South African youths. The analysis in this paper suggests a strong correlation between home background and the educational outcomes of the sample at hand. Broadly, the co-residence of biological parents in the household and the education attained by parents are found to be positively associated with educational outcomes of children. A number of variations on this theme are explored in the analysis. One critical limitation of the study at hand is the endogeneity which invariably plagues analyses of this kind. It must therefore be noted that while clear associations between home background and schooling outcomes emerge in the study, the analysis does not extend to establishing whether these associations are in fact causal. The clear associations that are uncovered certainly suggest that future research into the exact causal mechanisms behind these associations is likely to deliver important insights into the determinants of schooling outcomes in South Africa.

2. Home background and educational outcomes

There are a number of theoretical models which attempt to account for the observed association between family background and children's educational outcomes. It is important to distinguish between the various factors that are all encompassed under the umbrella term 'family background'. The term here refers an array of characteristics of a child's home

environment which have been found to be correlated with her educational outcomes, namely the composition of the household in which a child resides (the presence or absence of biological parents, the number of household members, the various relationships between the household members, et cetera) as well as specific characteristics of a child's parents. Of primary concern here are the various chains of causality that result in home background characteristics being associated with educational outcomes, that is, the transmission mechanisms that underpin this association. Although such chains of causality are notoriously difficult to isolate, this section discusses the various possible transmission mechanisms that are identified in the literature.

2.1. Household composition

The composition of the household in which a child resides has been shown to be associated with the educational outcomes she achieves. We know from a number of studies (such as Zoch (2015), Anderson (2005), Anderson et al (2001), among others) that children raised in homes where their biological mother is present achieve better educational outcomes than their counterparts who reside in mother-absent households. Similarly, children living with neither parent are disadvantaged in terms of educational outcomes (Heard, 2007: 435), whereas children living with both parents seem to have some educational advantage over those residing only with one parent (Townsend et al, 2002: 215). Children raised by two biological, married parents score higher on both mathematics and reading tests than children in other family types (Heaton et al, 2014: 104).

i) Parent co-residence

One broad set of models which attempt to explain such associations between child outcomes and household composition sees the outcomes achieved by children as a result of the investments children receive over the course of their childhood. According to these models, children who receive less investment suffer a range of negative outcomes compared to their peers, including poorer educational performance (Anderson, 2005: 1). In terms of an adult's decision to invest in a child, these models see humans as being primarily motivated by genetic fitness. This hypothesis is situated within a broader framework of human behaviour which relies on evolutionary biology. Evolutionary biology sees much of human behaviour as being motivated by the desire to pass on and protect one's genetic material. This is termed 'inclusive fitness', and was first introduced by Hamilton in 1964 (Anderson, 2005: 2). Also known as Hamilton's rule, this theoretical model and variants thereof see seemingly altruistic behaviour, such as a parent investing in a child, as the result of parents ensuring that their

genes survive. Thus, altruistic investment in children increases as the degree of relatedness between adult and child increases. Parents, being the closest relatives a child can have (apart from siblings) are therefore expected to invest the most in children, grandparents slightly less, more distant relatives even less and so forth.

Such models of inclusive fitness, or kin selection, thus explain the observed association between household composition and children's educational outcomes as the result of different levels of investment children receive, which in turn are determined by the degree of relatedness between children and their co-resident adults. All else being equal, children residing with both biological parents are therefore expected to achieve the best outcomes, those residing with only one parent are expected to achieve slightly worse outcomes, and those residing in households where neither parent is present worse outcomes still. Similarly, residing in a grandparent's household is expected to result in better outcomes for the child than, say, an uncle or aunt's household, or the household of an adult who is not related to the child.

According to these models, co-residence is considered important since a child spends the most time with adults who reside in her household. In a way, then, co-residence is assumed to be a relatively good measure of parental involvement. This assumption is generally problematic, but especially so in the South African context. Madhavan et al (2014: 7) report that in a sample of rural South African children more than 70% of respondents had social connections to their fathers, despite the fact that only 31% of similar-age children are reported to live with their fathers by Statistics South Africa. They further report that not only are non-co-resident fathers able to maintain contact with their children but many also provide financial support (Madhavan et al, 2014: 7). These findings suggest that social ties between father and child extend beyond co-residence for many South African children.

It is thus likely that father involvement is not well captured by father co-residence. It may well be that the same is true for mother involvement, though a lack of studies investigating whether absent mothers remain involved in their children's lives in South Africa makes it difficult to draw such conclusions at present. One of the aims of the present study is to investigate whether there may be some support for this notion. In other words, one of the things investigated in the multivariate analysis which follows is whether co-residence is in fact associated with children's educational outcomes. This might provide some indication of the value of parental co-residence as a measure of parental involvement.

ii) Intra-household resource allocation

The next factor that appears to be important for the investment children receive is the way in which decisions about resource allocation are made in the household. Such decision-making dynamics are influenced chiefly by different levels of bargaining power enjoyed by adults residing in the household. The study of intra-household resource allocation is relatively well-established in economics (Seebens & Sauer, 2007: 896). It seems clear from the empirical evidence that changes in an individual's relative intra-household bargaining power causes shifts in household expenditure patterns (Seebens & Sauer, 2007: 896). The degree of relatedness between adults in the household and children residing there is therefore not the only factor determining investment received by children. The latter is also affected by the relative bargaining power of co-resident adults. This bargaining power, in turn, is influenced by a number of characteristics of co-resident adults, such as their gender and age, as well as the relationship between them.

What is more, relative bargaining power of adults differs according to cultural norms regarding the appropriate role of different family members. Van der Vleuten (2016: 52) writes, for example, that family systems, that is, the "set of norms, common practices, and associated sanctions through which kinship and the rights and obligations of particular kin relationships are defined" differ markedly according to different cultural norms. In cultures where the nuclear family is considered the ideal living arrangement, parents are considered the highest authority on decisions regarding their children (Van der Vleuten, 2016: 55). In cultures where extended families (comprising three generations and/or relatives such as grandparents, aunts, uncles or married siblings who live together in one household) are considered ideal, decision-making involves more relatives than is the case with the nuclear family (Van der Vleuten, 2016: 55). This includes decisions over children in the household. Parents might enjoy less control over investment in their children in such families.

The relative bargaining power of co-resident adults also differs when living arrangements depart from what is considered 'ideal' in a specific cultural context, for example when parents separate or children are born to single mothers. Non-co-habitation by parents is particularly common in South Africa, partly as a result of the institutionalisation of labour migration under the apartheid regime (Madhavan et al, 2014: 5). As Madhavan et al (2014: 5) write, one of the legacies of apartheid is that "households function as 'stretched' residential units, with family members 'dispersed' between different households for reasons of work, education, care, support and housing." These diverse living arrangements clearly have important

implications for decision-making dynamics in the household which depart from 'traditional' models where relatedness between the adult and child is the most important factor determining investment in children. The diverse family structures observable in South Africa make the present study particularly relevant for studying the educational outcomes of children from a diverse range of households.

2.2. Parent characteristics

It is not only the presence of parents in the household which matters for the outcomes their children achieve. Specific characteristics of parents have also been shown to be highly correlated with children's educational performance. This idea has a lot of prima facie plausibility: few would contest that parents' traits are extremely important in influencing the educational outcomes of children. We know from the empirical literature that parents' levels of education are highly correlated with those of their children (Bowles & Gintis, 2002: 10). There is also a high correlation between the occupational status of parents and children across time (Bowles, Gintis & Osborne-Groves, 2005). Parents' education is also highly correlated with the cognitive development of a child: Noble et al (2012: 516) report that by the time of school entry, children from lower socioeconomic status (SES) backgrounds (which are characterised partly by lower levels of parental education) typically score between one-half and one full standard deviation below other children on most academic achievement tests.

It is clearly empirically established that characteristics of parents matter for the outcomes their children achieve. What is less clearly established in the literature is the exact mechanisms whereby a parent influences their child's outcomes. Are genetics at the heart of the correlation between parents' and children's outcomes? Or is it the different environments that parents of different educational and occupational statuses create for their children that sets them up for similar positions in life? Do parents influence their children's educational outcomes directly by helping them with homework and instilling in them behaviours that are rewarded in formal education? These questions are at the heart of a body of literature which seeks to untangle the different transmission mechanisms whereby parental characteristics specifically affect children's educational outcomes. The main findings of this literature are discussed below.

i) Genetics

The role of genetics in determining the influence parents have on their children's educational outcomes is the subject of a large body of literature. Specifically, the genetic inheritance of cognitive skill has received widespread attention (Bowles & Gintis, 2002: 10). Bowles and Gintis (2002: 10) report correlations of IQ between parents and offspring range from 0.42 to 0.72, indicating that a large part of a child's IQ is inherited from her parents. This does not imply, however, that genetic inheritance is necessarily a transmission mechanism whereby parents influence the educational outcomes of their children. For the latter to be true, it must first be established that IQ is in fact an important determinant of children's educational outcomes.

Heckman (2008) as well as Bowles & Gintis (2002), among others, argue that IQ is a much less important determinant of educational performance than is often assumed. They ascertain that performance in school is determined at least as much, if not more, by "non-cognitive" skills such as perseverance, self-confidence, motivation, discipline and time preference for leisure, as cognitive skills. According to these authors, the genetic inheritance of IQ is a relatively small part of the observed association between parent-child educational outcomes.

Although the genetic inheritance of IQ explains relatively little of the association between parent characteristics and child educational outcomes, this does not mean that the genetic inheritance of other personality traits is not important in causing the association between parent characteristics and children's outcomes. As Bowles and Gintis (2002: 13) argue, "the remarkable similarity between the incomes of identical twins compared with fraternal twins suggests that genetic effects may be important." Although these authors refer specifically to incomes, the high correlation observed between the educational attainment of parents and their offspring suggests that educational outcomes, too, have a genetic component. The work of Heckman and others regarding the importance of non-cognitive skills in educational performance suggests that this genetic component takes the form of IQ as well as other personality traits such as motivation and even tastes for education.

The exact magnitude of the effect of genetic inheritance in determining educational outcomes remains a controversial issue. The "nature versus nurture" debate has yet to reach a definitive conclusion, partly due to the difficulty in isolating genetic traits from the environment in which they manifest. As Heckman (2008: 307) explains, gene expression is governed by environmental conditions. This makes the expression of genes susceptible to environmental changes. It is for this reason that Heckman argues the "nature versus nurture" distinction is

obsolete: the traits we observe in individuals are created through the interaction of genes and environment, and to try and isolate either of these is a futile exercise. Nonetheless, it might be especially relevant in the South African context, where child fostering is common, to know how large a role genetics play in child educational outcomes. This might help disentangle some of the various transmission mechanisms that underpin the association between parental characteristics and child educational outcomes. Indeed, the South African context may provide an opportunity to isolate genetic effects from environmental ones, precisely because so many South African children are raised by adults other than their biological parents.

ii) Material resources

Parents may also influence the educational outcomes of children through the material circumstances they create for their families (Buis, 2012: 3). This is an especially important transmission mechanism in South Africa, where the quality of education received by children is clearly associated with their SES (Van der Berg, 2007). Aside from determining the quality of education children receive, the material circumstances of the family are likely to affect children's educational outcomes through the ability to pay for costs associated with education, such as school fees, uniforms, textbooks and stationery (Buis, 2012: 3). Economic resources may also affect the willingness of children to pay the opportunity costs of education, such as entering the labour market in order to augment household resources. Material resources may also be helpful in creating an environment that is conducive to success in school, such as a house that is big enough for all children to have a quiet space to complete their homework (Buis, 2012: 3).

The use of household income as a measure of family disadvantage is certainly not without its problems. Heckman (2008: 317) argues, for example, that the traditional focus on household income as a source of child disadvantage is probably misleading, since "affluent families may create impoverished childrearing environments" in much the same way that "[e]conomically disadvantaged families may provide ideal parenting environments." Nonetheless, one would be hard pressed to argue that a family's material resources do not constitute a very real constraint on its capacity to create an environment that is conducive to optimal childhood development. Household income therefore remains an important, albeit imperfect, mechanism whereby parents influence the educational outcomes of their children.

iii) Help with schoolwork

A third channel through which parents may influence children's education may be through their ability to help children with schoolwork. Parents with more education are expected to be more effective at helping their child succeed in school, by for example augmenting their child's ability to cope with the demands of the curriculum (Buis, 2012: 4). More educated parents may also be more effective in influencing their children about school-related decisions, such as which subjects to enrol in and when to leave school (Buis, 2012: 4). Learners, especially at older ages, are expected to be more likely to heed the advice of parents who have themselves been through the schooling system.

iv) Cultural capital

Buis (2012: 4) explains that an important mechanism whereby parents affect the educational outcomes of their children is through the transmission of what he calls 'cultural capital'. He explains that

> "…more cultural capital means that a child tends to have the language and dialect, cultural preferences, and ways of interacting with others that are viewed positively by teachers. This will influence the performance of children, both because children will be less likely to view school as a hostile environment, and because these characteristics are positively sanctioned by teachers."
>
> - Buis (2012: 4)

Buis is thus not referring to 'culture' as is often understood in the South African context, as a language group with a shared heritage. Rather, he is referring specifically to the behaviours that are rewarded in the school environment. While it falls beyond the scope of this paper to determine whether this *should* be the case, there is little doubt that certain behaviours are rewarded by teachers, and consequently learners exhibiting those behaviours are likely to succeed in school. Parents in possession of the type of cultural capital that is rewarded in formal education are likely to instil this in their children, thus constituting a further mechanism whereby parents influence the educational outcomes of their children.

v) Occupation

There are three mechanisms which have been put forward as potential channels whereby parents' occupation affects children's educational outcomes. These are through material resources earned, through the stress associated with an occupation, and through the creation of non-cognitive skills in children. Firstly, parents' occupation is important for children's

educational outcomes in that it largely determines the material circumstances in which children grow up (Buis, 2012: 5). The ways in which material resources impact on children's education have already been discussed above. Secondly, Buis (2012: 5) postulates that the stress associated with an occupation is likely to impact on a parent's ability to invest time and energy in her children. Thirdly, Bowles and Gintis (2002: 21) write that children are likely to be influenced by parents' experiences in the workplace. Specifically, they argue that these experiences are pivotal in shaping the non-cognitive skills parents instil in their children. These authors cite evidence of parents who experience high degrees of autonomy at work encouraging curiosity, self-control, happiness and independence as values for their children, while those who are closely supervised at work emphasise conformity to external authority. The former non-cognitive skills are thought to impact positively on a child's performance in school, while the latter do not share the same positive association with performance.

2.3. Which parent?

All of these hypotheses about the mechanisms by which parents affect their children's educational outcomes have implications for, firstly, *whether* these mechanisms should differ depending on the gender of the parent, and secondly, if these mechanisms do differ by the gender of the parent, *how* these differences manifest.

Neither the inclusive fitness nor the genetics hypotheses imply a difference in the mechanism whereby a parent influences their child's educational outcomes based on the gender of the parent. Since a child receives 50% of her genes from either parent, according to the genetics hypothesis, both mother and father have equal influence over the genes they pass on to their child. Similarly, since mother and father share the same degree of relatedness with a child, according to inclusive fitness both should invest equally in their child.

The mechanisms whereby parents aid their children's progress through school by helping them with homework and influencing their education-related decisions, as well as whereby parents transmit cultural capital to their children, are likely to be more effective among parents who spend more time with their children. Thus we might expect the characteristics of the parent who spends the most time their child to have the largest impact on the child's educational outcomes. Since it is traditionally the mother who spends the most time with children, one might expect characteristics of the mother to have a greater impact on her children's schooling outcomes than characteristics of the father, however this effect is likely to be smaller in households where the mother also works (Buis, 2012: 6).

According to another broad set of hypotheses, the parent that has the largest influence on the child's educational outcomes is determined not by which parent spends the most time with the child, but which parent is 'dominant', in the sense that they have more education or enjoy a higher occupational status, than the other. This set of hypotheses is broadly termed the 'dominance model' (Buis, 2012: 6). The implication of this model is that it does not matter whether it is a child's mother or father who is well educated, for example. As long as the child has one parent who is educated, their educational outcomes are likely to be positively influenced by that parent.

A third set of hypotheses sees neither the time spent with a parent nor the higher educational or occupational status enjoyed by a parent as the mechanisms whereby parents influence their children's education. Rather, this model postulates that sex-roles are most important in transmitting characteristics from parent to child. That is, daughters are assumed to be "primarily oriented towards their mother and sons to their father because the same-sex parent is perceived by children to have more relevant information for their situation" (Buis, 2012: 6). Buis (2012:6) explains that it is useful to distinguish between a strong and weak version of this hypothesis: the strong version would be that it is only the occupation or education of the same-sex parent that matters, while a weak version would assert that both parents matter, but the occupation or education of the same-sex parent matters most.

2.4. Parental involvement

Apart from the implications for the relative influence of mothers compared to fathers on children's outcomes, the hypothesis that sees the amount of time parents spend with their children as important in determining the effectiveness of certain parent-child transmission mechanisms has important implications for our understanding of the exact mechanisms whereby parents influence the educational outcomes of their children. Mechanisms like genetics should work even when parents are not present in the household: a child's genetic make-up is determined at conception, thus one might therefore expect this mechanism to work even when parents are absent from the households where their children reside.

By contrast, the other transmission mechanisms discussed above, such as the influence parents have on their children's educational decisions, as well as their effectiveness in helping their children succeed in school, are likely to be influenced by the amount of time the parent spends with the child. These mechanisms are therefore expected to be more effective in households where parents live with their children – after all, children spend the most time

with those adults who live with them. In South Africa, where many children live apart from their mother, and the vast majority of children do not reside with their fathers, one might expect many of the mechanisms whereby parents pass on their characteristics to their children to be relatively ineffective at best, and non-existent at worst.

Following on from this literature, the remainder of this paper sets out to examine associations between a number of family background characteristics and the educational outcomes of a sample of South African youths.

3. Methodology

3.1. The sample

The present study makes use of a sample of school-age individuals captured in the National Income Dynamics Study (NIDS). NIDS is a nationally representative survey, conducted in 2008, 2010, 2014 and 2016. Here the four waves of NIDS are used as a cross-sectional dataset, that is I use data from all four rounds, but do not consider respondents who appear more than once across the eight years, taking the most recent year that a respondent was interviewed as the year of observation.

I further restrict my sample to respondents between the ages of 15 and 17. This was done to get the most meaningful metric of educational outcomes possible in the NIDS data. Unfortunately, the NIDS data does not contain very detailed information about respondents' educational outcomes. The survey only contains information about whether a respondent is currently enrolled in education, the number of years of education they have completed, and the highest tertiary qualification they have completed, if any. Although the first wave of NIDS does contain detailed information about grade repetition – whether a respondent has ever completed a grade, as well as which grades they have completed – the fact that the other waves do not contain this information meant that this information could not be used in the present analysis. It is because of this limitation that the sample is limited to those respondents between the ages of 15 and 17. Van Wyk (2015: 8) shows that in South Africa, grade repetition only becomes widespread in Grade 9, and learners begin to drop out of school in large numbers around Grade 11. Before Grade 9, that is, before age 15, neither grade repetition nor dropout is common. Given that educational attainment and enrolment are the only education variables available in the NIDS data, including respondents younger than 15 therefore does not yield very meaningful results. In terms of the upper bound of the chosen sample, 17 was chosen since I am concerned with the educational outcomes of school-age

respondents. Although learners are supposed to turn 18 in matric, when including 18-year-olds in the sample one encounters the problem that many respondents have already completed secondary school and therefore cannot really be considered 'school-age'. One might overcome this problem by only including 18-year-olds who are still enrolled in secondary school, however this eliminates respondents who are 18 but have dropped out of school from the sample. Given the large scale of dropout specifically at this age, removing dropouts from the sample not only results in a very large reduction in sample size, but also creates a biased sample consisting only of those who have not dropped out of school. Since this study is concerned with the educational outcomes of all South African youths, not just those who have stayed in school, it was decided to simply only include 15 to 17-year-olds in order to overcome these problems. This provides a nationally representative sample of 7 374 respondents.

3.2. Home background variables

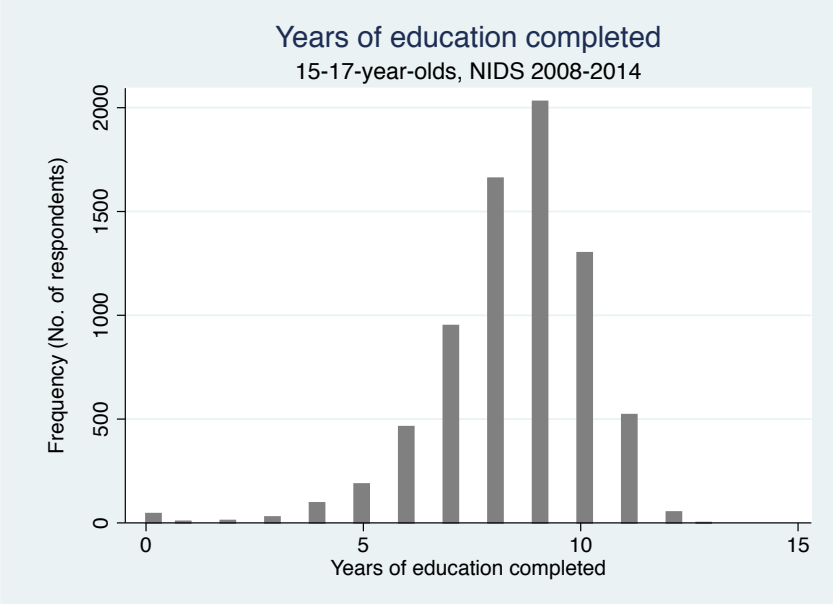
The NIDS data contains a host of variables relating to home background, such as household size, the gender of the household head, the respondent's relationship to their household head, the relationship between the respondent's parents, the years of education the respondent's parents have completed, their parents' occupation, and household income. Further one can link relatives who reside in different households by using the unique codes assigned to each respondent. One can also use this unique identifier to derive whether a child's parents are present in the household. Together this information can provide a detailed description of the household circumstances of each respondent.

3.3. Outcome variable: educational attainment

The outcome variable in the present study is the years of education that a respondent has attained. Although a relatively crude measure of educational performance, attainment nonetheless provides a good indication of respondents' educational outcomes, especially in this age group. The standard deviation of this variable is 1.72, around a mean of 8.41, indicating a relative amount of variation. Notably, the mean is quite a bit lower than one might expect for this age group. This is especially startling if one looks at the age distribution of the sample: about 24%, 38% and 38% of respondents are 15, 16 and 17 years old, respectively. Since 15-year-olds are supposed to have completed at least grade 8, 16-year-olds grade 9, and 17-year-olds grade 10 in order to be considered 'on track', a mean attainment of 8.41 years of education for this age group is lower than should be the case, had all the respondents completed the expected grade for their age. This low mean thus suggests that

dropout or grade repetition, or both, are quite prevalent among respondents in the given sample.

Figure 1: Educational attainment of respondents



This is further evident if one looks at the distribution of educational attainment of the sample (Figure 1). While the median years of education completed is grade 9, there are startlingly many respondents whom have not yet completed primary schooling (roughly 1700 out of 7374 have completed grade 7 or less). This means more than 20% of the sample is at the very least one year behind where they should be in terms of grade attainment - but the true proportion of overage learners is likely much larger. The fact that there is such variation in the distribution of grade attainment among the given sample makes the years of schooling attained, i.e. the highest grade completed, a useful measure of educational performance for this group of respondents.

3.4. Method

First, I present some descriptive statistics of the home background variables mentioned above. This is done in order to sketch a broad picture of the home background circumstances of the given sample of 15-17-year-old South African youths. Simple OLS regressions are then used to estimate correlations between these family background variables and the years of education attained by respondents. Following this, I include interaction terms in order to determine

whether the estimated correlations differ when key explanatory variables are interacted with other variables. The results of this analysis are presented below.

4. Results

4.1. Descriptive statistics

4.1.1. Household income

In terms of household income, it is useful to consider proportions of respondents who fall into quintiles of per capita household income. These quintiles are created by arranging the per capita household incomes of all respondents in the original NIDS dataset from smallest to largest and dividing the whole sample into fifths. The distribution of respondents in the current sample across income quintiles is as follows: 29% of the sample come from quintile 1 households, a further 29% from quintile 2 households, 22% in quintile 3 households, 13% in quintile 4 households, with only the remaining 7% residing in quintile 1 households. The fact that respondents are not equally distributed across quintiles indicates that poorer households have more members between the ages of 15 and 17 than better-off households. Specifically, the large differences in proportions of respondents in different income quintiles indicates that the differences in the number of 15-17-year-old respondents between the income quintiles are substantial.

4.1.2. Household composition

Table 1 below shows the proportions of respondents to whom different characteristics related to household composition can be attributed. The table shows that the overwhelming majority of learners reside in households headed by women (64%). Posel (2001: 658) finds that women are very unlikely to be identified as the heads of household when they reside with their spouses: in her sample of South African households, only 3% of females who were identified as the heads of household were living with their spouses. From this we can infer the households identified as being female-headed in current sample are likely to be households where the head is also not living with a male spouse. In other words, it is likely that the majority of respondents in the sample reside in households headed by a single woman.

Table 1: Proportions of respondents residing in different types of households

Variable	Proportion
Household is female-headed	64%

Three-generation household	12%
Skip-generation household	20%
Mother is resident in the household	63%
Father is resident in the household	24%
Both parents are resident in the household	24%

The proportion of respondents who reside in three-generation households, that is, households where a parent and grandparent are both resident, is relatively high, at 12 percent. Skip-generation households, where a child resides with a grandparent in the absence of a parent, are also relatively common, with 20 percent of respondents residing in such homes.

The proportion of respondents residing in households where their mother is resident is worryingly low, at only 63 percent. This is much lower than even the western and southern African average of 75-84 percent of children residing in mother-present households, which in turn is considered low by international comparison (Zimmerman, 2002: 558). Even more worrying is the proportion of respondents whose fathers are absent from their households, at just about three-quarters. The proportion of respondents who reside with both parents is exactly equal to that of respondents who reside with their fathers, suggesting single father households are extremely uncommon. A quick check of the number of learners residing in households where their father is resident but not their mother reveals this is in fact the case: only 14 out of the sample of 7 374 15-17-year-olds reside in single-father households. Clearly, single-parenthood is overwhelmingly skewed towards women in the given sample.

4.1.3. Parents' education

Figure 2: Parents' education

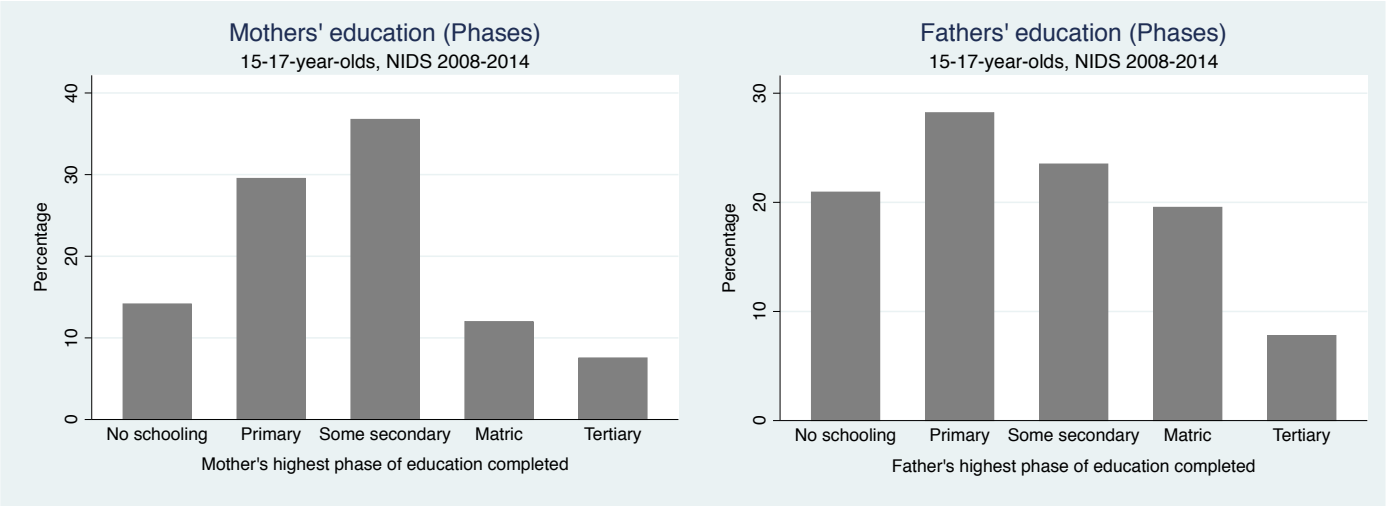


Figure 2 shows the proportions of respondents whose parents (co-resident or absent) have completed various phases of education. It must be noted that there are likely some reporting issues relating to these parental education variables. The question about parents' education is asked at the level of the individual, simply as "What is the highest grade in school that your mother/father has successfully completed?" This information was combined with that from the question "Did your mother/father successfully complete any diplomas, certificates, degrees outside of school? If yes, what is the highest level of education your mother/father has completed?" to create a variable reflecting respondents' parents' highest level of education. Potential reporting issues are likely since children may not know the exact school grade or tertiary qualification their parents have completed. This may especially be true in cases where the parent is absent from the household where the child resides. Given that a third of respondents in the current sample reside in mother-absent households, and three-quarters reside in father-absent households, it is likely that inaccurate reporting of parental education, especially father's education, is widespread.

Parents' education was grouped into 'phases' of education in an attempt to somewhat correct for this. Children may not know whether their parents completed, for example, Grade 4 or Grade 5, a problem which is solved when grouping all primary school grades together as either having completed primary school (Grade 7) or not. A more serious concern than small estimate errors is that many respondents may not know their absent parent's level of education at all. This is more likely a concern for paternal education given the large number of absent fathers in the sample. It may even be that respondents with absent fathers may not

know their father, in which case any information about their father's education would be pure guesswork. What is more, it is likely that children overestimate their father's level of education in such circumstances. It is therefore probable that, on average, the levels of paternal education reported here are inflated. It is important to keep this in mind in the analysis that follows.

It is noteworthy that despite the fact that parental education, especially father's education, is likely overestimated in this sample, the proportion of respondents whose fathers have no education is still large, at roughly 21%. It must be kept in mind, however, that 'no education' here means the respondent reported their father as not having completed primary school. Nonetheless, this is a high proportion, especially when one considers that the true proportion is likely to be even larger.

It is clear that the proportion of respondents whose fathers have no education is larger than that whose mothers have no education (roughly 21% compared with 14%). The proportion of respondents whose fathers have some secondary education is also smaller than that of mothers: while only about 23% of fathers have some secondary education, this proportion is around 38% for mothers. However, the opposite is true for the completion of matric: the proportion of respondents whose fathers have completed matric is much larger (just under 20%) than those whose mothers have completed matric (about 11%). Once again, the overestimation issue may be important here. The proportions of learners whose mothers and fathers have attained some tertiary education is, albeit very small, roughly equal.

4.1.4. Parents' occupation

The manner in which parents' occupation is recorded in the NIDS data makes it difficult to assign a scale to this variable. Occupations are coded only in broad terms, such as "service and sales workers" or "crafts and related trade workers". It is conceptually problematic to assign a higher "value" to one of these occupations over another. For this reason, it was decided to use only information about whether a parent has ever worked as an indication of their workplace experiences. While it is difficult to assign specific work experiences to a broad occupational class such as 'professionals' versus 'service and sales workers', for example, it is more straightforward to compare parents who have been chronically unemployed, i.e. who have never worked, with parents who have held a job at some point. 34% of the sample at hand have mothers who have never worked, while 32% have fathers who have never worked. About half of those whose mothers have never worked also have fathers who never worked.

In other words, roughly 16% of the sample have two parents who have never worked. If parents' experience of chronic unemployment impacts negatively on their child's schooling outcomes, we might therefore expect a relatively large proportion of respondents to be disadvantaged in terms of this aspect of home background.

4.2. Multivariate analysis

The most basic specification, Model 1, does not include controls for any family background variables. The coefficient on age has the expected size and significance. The large positive and significant coefficient on the female dummy variable is noteworthy if not altogether surprising: girls in the sample have attained just about two-thirds of a year more education, on average, than boys, keeping age, race and household income constant. This supports findings elsewhere in the literature that boys in this age group are much more likely than girls to repeat a grade or drop out of school altogether (Van Wyk, 2015).

The model further controls for the effect of race on educational attainment by including dummy variables for white, Indian and coloured respondents. Black respondents therefore constitute the reference category. It is quite surprising that the coefficient on the white dummy is negative albeit insignificant – one would certainly expect white respondents to perform better in terms of grade attainment given the vastly unequal educational outcomes of black and white learners generally reported in South Africa (Spaull, 2012). One explanation for this might simply be that the sample of white respondents is too small to precisely estimate effects for this group: barely 1.5% of the sample at hand is made up of white respondents. Given the imprecision with which educational outcomes are estimated for white respondents, this coefficient likely does not reflect the true effect of being white on grade attainment. Controlling for race in all the models ensures distortions created by this imprecise estimation are eliminated.

Per capita household income (expressed in thousands of Rand), as well as per capita household income (again in thousands of Rand) squared, are also included in Model 1. The coefficients on these variables are of the expected size and significance. The negative coefficient on per capita household income per thousand indicates decreasing returns to income, in other words more income is associated with smaller increases in educational attainment at higher levels of household income.

Table 2: Household composition regressions

	Years of education (1)	Years of education (2)	Years of education (3)	Years of education (4)
Age	0.666 (27.75)**	0.668 (27.85)**	0.668 (27.82)**	0.668 (27.86)**
Female	0.657 (17.65)**	0.657 (17.67)**	0.656 (17.62)**	0.775 (10.56)**
White	-0.052 (0.31)	-0.056 (0.34)	-0.046 (0.27)	-0.036 (0.21)
Coloured	-0.029 (0.51)	-0.051 (0.89)	-0.045 (0.78)	-0.043 (0.75)
Indian	0.678 (3.30)**	0.633 (3.08)**	0.637 (3.10)**	0.639 (3.11)**
Income (thousands of Rand)	0.211 (12.30)**	0.199 (11.18)**	0.201 (11.26)**	0.200 (11.21)**
Income^2 (thousands of Rand)	-0.000 (8.36)**	-0.000 (7.83)**	-0.000 (7.88)**	-0.000 (7.82)**
Mother resident in household		0.082 (1.93)	0.072 (1.69)	0.172 (2.87)**
Father resident in household		0.110 (2.29)*	0.147 (2.79)**	0.193 (2.71)**
Number of children in household		-0.015 (1.55)	-0.016 (1.55)	-0.023 (1.58)
Household female-headed			0.073	0.071
[Mother in household]*Female				-0.200 (2.37)*
[Father in household]*Female				-0.092 (0.97)
Female*[Number of children]				0.013 (0.69)
Constant	-2.883 (7.40)**	-2.945 (7.52)**	-2.985 (7.61)**	-3.056 (7.76)**
R^2	0.15	0.15	0.15	0.15
N	7,349	7,349	7,349	7,349

* $p<0.05$; ** $p<0.01$

The second model includes controls for various aspects of household composition (column 3, Table 2). 'Number of children' in the household refers to a variable created from the NIDS household roster which indicates the number of children below the age of 15 who are resident in the household. It is somewhat surprising that the coefficient on this variable is also insignificant – one might expect that more young children in the household may impact negatively on respondents' educational outcomes through draining more of the household's resources, or through the respondents having to help care for the younger children. It seems that neither such effect exists.

The coefficients on per capita household income remain roughly similar in the second model. The effect of household income on educational attainment therefore remains roughly the same when one includes controls for various aspects of household composition. This echoes findings elsewhere in the literature of the importance of household material resources in determining the quality of education received by learners in South Africa (Spaull, 2012).

The coefficient on the dummy variable indicating whether the respondent's mother is present in the household is surprising. As discussed in the literature review, a mother's presence in the household is found to be positively associated with children's schooling in a number of studies. It is odd that the coefficient here is not statistically significant, suggesting that respondents who live with their mothers are statistically no different in terms of years of schooling attained, on average, than learners whose mothers are absent from their households.

While the results suggest that a mother's co-residence in the household is not statistically significantly associated with respondents' educational attainment, they indicate that such an association does exist for father co-residence. Respondents who live with their fathers have attained 0.11 more education than their counterparts who do not live with their fathers. It is important to keep in mind the descriptive statistics on mother and father co-residence when interpreting this coefficient. Only 14 respondents lived with only their father and not their mother. In other words, just about all the respondents who lived with their fathers also lived with their mothers. This means father co-residence is really a measure of both parents being co-resident in the household. The coefficient on father co-residence therefore essentially reflects the effect of living with both parents. The same is not true for mother's presence, since 61% of mother-present households are also single mother households. The coefficients on mother and father co-residence therefore reflect the effect of living with a single mother and

that of living with both parents, respectively. It is therefore not really the co-residence of a mother that is insignificant, but rather living with a single mother compared with living in a mother-absent household. Similarly, respondents who live with both parents have attained more education, on average, than their counterparts who live in households where neither biological parent is present.

A Wald test of significance was conducted to determine whether the coefficient on mother co-residence is significantly different from that of father co-residence. A p-value of 0.701 suggests the coefficients are not statistically significantly different from each other. In other words, the effect captured by father co-residence is not significantly different from that of mother co-residence. These two variables capture the same effect, the effect of father co-residence being larger and therefore appearing as significant when mother co-residence does not. This finding suggests mother co-residence does have a positive association with educational attainment, which only becomes significant when the father is also co-resident in the household.

Model 3 contains the same household structure controls as Model 2, with a dummy variable indicating whether the household is headed by a woman. The coefficient on this variable is not significant, indicating that respondents residing in such households are not statistically significantly different from respondents residing in male-headed households. This is interesting since findings elsewhere in the literature suggest female-headed households may be considered 'vulnerable' (Posel, 2001: 664) in terms of a number of dimensions, and that children who reside in female-headed households may suffer from lower educational outcomes (Handa, 1994: 1535). It would appear that this vulnerability, if it exists, is not notably associated with the schooling attainment of the sample of respondents at hand. It must be noted, however, that is likely that the coefficient on female headship is insignificant since the effect of female headship may already be captured by the mother co-residence variable. A Wald test of significance testing whether the effect of female headship differs from that of mother co-residence has a p-value of 0.980, indicating that these two variables measure almost precisely the same effect. This indicates something important about household structure, namely that even in father-absent households where the mother is not identified as the household head, the household is still headed by a woman. In other words, when the household is not headed by the father, the household is either headed by the mother or another woman. Male heads who are not the father of the child in question are extremely rare in this sample. This has important implications for the interpretation of the coefficients in both Models 2 and 3. Since the Wald test of significance shows female headship

and mother co-residence (in the absence of the father) measure essentially the same thing, the interpretation of the coefficient on mother co-residence produced by Model 2 now changes: it is not mother co-residence per se which is not significantly associated with schooling attainment, but rather female headship. This in itself is interesting, since it supports the notion that the vulnerability typically associated with female headship is not picked up here. Furthermore, the insignificant coefficient on mother co-residence might be due to the fact that this variable in fact captures the net effect of two opposing effects. On the one hand, we expect mother co-residence to have a positive effect on educational attainment, and on the other, female headship is expected to impact negatively on a child's schooling. The fact that these two effects oppose each other could result in an insignificant net effect. While it is not possible to confirm this hypothesis with certainty, findings elsewhere in the literature of the positive impact of mother co-residence on schooling outcomes as well as the vulnerability associated with female headship, together with the Wald test result that these two variables measure essentially the same effect in the given sample, all support the notion that mother co-residence does have a positive effect, but that this is countered by the negative effect of female headship.

The coefficient on father co-residence, which essentially measures the effect of living with both parents, increases in size and significance when female headship is included. This suggests the effect of living with both parents is estimated more precisely when including a control for female headship, supporting the notion that the variable indicating father co-residence essentially captures the effect of male headship as well. It is therefore unclear whether it is the presence of both parents in the household which positively affects respondents' schooling outcomes, or whether the positive effect one observes is actually that of residing in a male-headed household. The coefficient on father co-residence is not statistically significantly different from that of mother co-residence.

Glick & Sahn (2000: 64) report that various elements of household structure may impact girls and boys differently. For example, they find that mothers may favour girls in their spending on education. They also report that girls are more likely to carry part of the childcare burden than boys, thus one might expect the educational outcomes of girls to be more likely to be negatively influenced by the number of children in the household. In order to determine whether such gender effects may be present for the current sample, the dummy indicating the presence of a mother in the household was interacted with the gender of the respondent. The same was done for the presence of the father in the household. The number of children in the

household was also interacted with the gender of the respondent. The results are reported in column 4 of Table 2.

A number of interesting results emerge. Firstly, the coefficient on the dummy indicating a mother's residence in the household increases, which the previous model showed essentially indicates the household is female-headed, from being insignificant in the previous models to 0.172, which is significant at the 95% level. This suggests that when allowing for the effect of a mother's co-residence to vary depending on the gender of the child, respondents who reside in households headed by their mother or in households headed by another woman where their mothers are also present have attained 0.172 years more education, on average, than their counterparts whose mothers are absent from their households.

The coefficient on father's co-residence also changes when including such interaction effects. In the third model, where no interaction effects are included, the coefficient on father's co-residence is 0.147 and significant at the 95% level. In the fourth, the coefficient increases in size to 0.193 and maintains the same level of significance. This suggests that when allowing for a different effect of father's co-residence depending on whether the respondent is a boy or girl, respondents whose fathers are resident in their households attained 0.193 years of education more, on average, than their counterparts whose fathers are absent from their households, and, by implication, who reside in households where neither biological parent is present.

The coefficients on the interaction terms themselves provide some indication of why an association between parent co-residence and attainment only emerges when one allows for this association to differ depending on the gender of the child. The coefficient on the interaction term between mother co-residence and the female dummy is -0.200. This suggests that on average, girls who reside in female-headed households where their mother is present attained 0.200 years *less* education than boys residing in such households. Since the coefficient on mother's co-residence is 0.172, this means the net effect of a female-headed household and mother's co-residence is -0.020 for girls, slightly reducing the female advantage. In other words, the positive association of mother co-residence vanishes for girls, and in fact girls in female-headed households where their mothers are co-resident have attained, on average, 0.020 years of education *less* than girls residing in mother-absent female-headed households.

This may further explain why there is no significant association between mother co-residence and attainment in the second model, but one is observed in the fourth model, in addition to the opposing effects of mother co-residence and female headship as discussed above. On average, mother co-residence in female-headed households has no positive association with attainment since mother co-residence in female-headed households has a positive association with attainment for boys but a slightly negative association with attainment for girls. The net association across the sample of girls and boys is therefore insignificant. Once one allows for gender differences, however, this average effect between boys and girls can be untangled and a positive association emerges between mother-present female headship for boys, while no such association exists for girls. This is an extremely interesting result, given findings elsewhere that seem to suggest if any gender differences exist for the association between mother co-residence and educational attainment, it is that girls' education is *more* strongly positively associated with mother co-residence than is the case for boys. Unfortunately, it is not possible to ascertain the exact extent to which these opposing gender effects result in the insignificant net effect of mother co-residence in female-headed households. In any event, given the fact that female headship is likely a large part of the explanation for this insignificant net effect, it is unlikely that the negative effect of being a girl is so large as to entirely counter the positive effect of mother co-residence in female-headed households.

The interaction term between father's co-residence and the female dummy is insignificant, suggesting that the association between father's co-residence and attainment does not differ depending on the gender of the respondent. In other words, the positive effect of living with both parents and/or living in a male-headed household is the same regardless of whether the respondent is a boy or a girl. Once again, the coefficients on mother and father co-residence respectively are not significantly different from each other.

The coefficient on the interaction term between the number of children in the household and the female dummy is also insignificant. The negative effect of other (usually younger) children in the household on the educational outcomes of girls specifically that is observed elsewhere in the literature is therefore not picked up here.

The fifth model specification (reported in Table 3 below) includes variables indicating parents' highest 'phase' of education attained – no education, primary school, some secondary schooling, matric, and some tertiary. Parental education was divided into these phases in order to get the most accurate reflection of the association between parental

education and attainment. It was thought that using years of parents' schooling might be problematic for two reasons.

Firstly, the accuracy of this variable is questionable, for reasons explained in the Descriptive Statistics section of this paper. Secondly, it is doubtful whether a single year of additional education reveals much about the characteristics of an adult, such as their intelligence, perseverance, self-discipline, etc. and that it rather reflects opportunities and restrictions on such opportunities for studying at the time they were young. This is especially true at higher levels of schooling, since completion up to lower grades is mandatory. Parental education was therefore divided into the aforementioned phases in order to reduce this noise around the parental education variables.

Table 3 presents the findings that result from adding variables indicating parental education to the second model specification. Four dummy variables for each parent's education are added, with 'no education' as the reference category. The first thing to note about this model is the reduction in the number of respondents from 7349 in the previous four models to 5764 in this one. This is most likely due to missing information about parents' education, especially non-co-resident fathers who constitute almost three quarters of the original sample. Although missing information can be dealt with in a number of ways, it was decided to simply drop these observations from the sample. Naturally doing so results in a somewhat biased sample, which only includes those respondents for whom information on parental education is known. It is important to keep this in mind in the discussion that follows.

The first interesting result that emerges from these specifications is the large decrease in the coefficient on income – the coefficient falls from around 0.2 in the previous models (which do not control for parental education) to 0.067 in the fifth model. This suggests that much of the association between household income and educational attainment is in fact due to the positive association between income and parental education. This would mean that it is not the effect of income alone which benefits children in terms of their educational attainment, but rather that income is closely related with parental education, and the latter explains a large part of the association between income and educational attainment. This constitute an interesting result, since it suggests it is not only material resources which impact positively on children's schooling outcomes – even when controlling for income, a large and significant effect of parental education remains.

Table 3: Parental education regressions

	Years of education (5)	Years of education (6)	Years of education (7)
Age	0.685 (25.96)**	0.684 (25.94)**	0.683 (25.93)**
Female	0.624 (15.43)**	0.827 (7.36)**	0.622 (15.39)**
White	-0.136 (0.79)	-0.101 (0.58)	-0.128 (0.74)
Coloured	-0.060 (0.94)	-0.066 (1.03)	-0.061 (0.95)
Indian	0.470 (2.32)*	0.466 (2.31)*	0.454 (2.25)*
Income (thousands of Rand)	0.067 (3.43)**	0.066 (3.37)**	0.081 (4.05)**
Income^2 (thousands of Rand)	-0.000 (2.99)**	-0.000 (3.00)**	-0.000 (3.44)**
Mother resident in household	0.056 (1.16)	0.060 (1.24)	0.392 (3.43)**
Mother completed primary schooling	0.176 (2.67)**	0.219 (2.33)*	0.462 (3.76)**
Mother completed some secondary schooling	0.486 (7.26)**	0.519 (5.47)**	0.758 (6.51)**
Mother completed matric or more	0.937 (12.11)**	1.167 (10.90)**	1.222 (9.73)**
Father resident in household	0.206 (4.19)**	0.207 (4.22)**	0.083 (0.88)
Father completed primary schooling	0.069 (1.15)	0.112 (1.32)	-0.010 (0.14)
Father completed some secondary schooling	0.293 (4.55)**	0.317 (3.48)**	0.200 (2.68)**
Father completed matric or more	0.418 (6.66)**	0.463 (5.20)**	0.419 (5.95)**
Constant	-3.676	-3.764	-3.861
R^2	0.19	0.19	0.19
N	5,764	5,764	5,764

* $p<0.05$; ** $p<0.01$

Table 3: Parental education regressions (cont.)

	Years of education (5)	Years of education (6)	Years of education (7)
[Mother completed primary schooling]*Female		-0.090	
		(0.69)	
[Mother completed some secondary schooling]*Female		-0.070	
		(0.52)	
[Mother completed matric or more]*Female		-0.463	
		(3.12)**	
[Father completed primary schooling]*Female		-0.088	
		(0.74)	
[Father completed some secondary schooling]*Female		-0.046	
		(0.36)	
[Father completed matric or more]*Female		-0.096	
		(0.78)	
[Mother resident in household]*[Mother completed primary schooling]			-0.395
			(2.77)**
[Mother resident in household]*[Mother completed some secondary schooling]			-0.388
			(2.86)**
[Mother resident in household]*[Mother completed matric or more]			-0.417
			(2.78)**
[Father resident in household]*[Father completed primary schooling]			0.224
			(1.82)
[Father resident in household]*[Father completed some secondary schooling]			0.289
			(2.21)*
[Father resident in household]*[Father completed matric or more]			-0.097
			(0.72)
Constant	-3.676	-3.764	-3.861
	(8.43)**	(8.58)**	(8.75)**
R^2	0.19	0.19	0.19
N	5,764	5,764	5,764

* $p<0.05$; ** $p<0.01$

Table 3 shows that the coefficients on all the maternal education variables are significant. Respondents whose mothers have completed primary school have themselves attained 0.176 years of education more, on average, than their counterparts whose mothers have not

completed primary school. Respondents whose mothers have completed some secondary schooling have attained 0.486 years of education more than their peers whose mothers have not completed primary school. Respondents whose mothers have completed matric or some tertiary education have themselves completed 0.937 grades more than their counterparts whose mothers have not completed primary school. This is an extremely important result, since it implies even when controlling for mother co-residence in female-headed households, respondents whose mothers have completed matric or more are almost a whole grade ahead of learners whose mothers have not completed primary school. The coefficients on father's education follow a similar pattern, although the coefficients on paternal education are roughly half the size of those on maternal education for almost all phases of education. An exception to this is having a father with primary school, the effect of which is not statistically significant compared to having a father who has not completed primary school. This is in contrast to having a mother with primary schooling, which does have a significant positive effect relative to having a mother who has not completed primary school for the sample on average.

Interestingly, the coefficients on mother and father co-residence are significantly different from each other in the fifth model. The effect of mother co-residence in female-headed households is therefore different from the effect of both parents residing in one's household. This is an interesting result, since these two effects are not significantly different from each other in the previous model specifications, where parental income is not controlled for. The fact that they do differ from each other in the fifth specification may suggest that the larger impact of the absence of the mothers than fathers is obscured in the previous model by the exclusion of controls for parental education.

An important concern is whether the effect of having a mother who has a certain level of education is significantly different from that of having a father who has completed a certain level of education. Assortative mating (where individuals choose mates who have certain characteristics in common with them, such as, for example, education) makes it likely that respondents have parents with similar levels of education. It is therefore important to distinguish whether there is a separate effect of maternal versus paternal education. A Wald test of the difference between a having a mother who has completed matric or more compared with a father who has done so results in a p-value of 0.000, indicating that having a mother with matric or more has a much larger effect on one's education from having a father with matric or more. The same test for having a mother who has completed some secondary

education and a father with some secondary education produces a p-value of 0.071, which suggests the two variables also measure different effects. A test shows that the effect of having a mother with primary schooling is not significantly different from that of having a father with primary schooling.

While it is clear that parental education is important for the educational attainment of the sample at hand, and that this positive effect is separate from the effect of income on attainment, the exact mechanism which causes this effect is less clear. As discussed in the literature review, more educated parents might impact positively on their children's ability to succeed in education by cultivating in them the skills (cognitive as well as non-cognitive) that are rewarded in formal education. Genetics might also be at the root of this association. Alternatively, more educated might have the knowledge and resources to send their children to better quality schools. To better untangle these effects, Model 6 explicitly interacts the parental education dummies with the parental co-residence dummies. The results from the sixth specification are discussed below.

As is the case in the fourth model, the sixth model includes interaction terms between parental education and the female dummy to establish whether the effects of parental education differ depending on the gender of the respondent. As the estimates in column 3 of Table 3 show, only one of the interaction terms between maternal education and the female dummy is significant, namely that between the latter and a mother having completed matric or more. This coefficient suggests that girls whose mothers have not completed matric or more have themselves attained 0.463 grades less than boys whose mothers have completed this level of education. Since the coefficient on having a mother with matric or more is very large, at 1.167, this negative effect of being a girl does not cancel out the large positive effect of having a mother with matric or more. Nonetheless, the negative coefficient on the interaction term once again presents a very puzzling result, namely that a mother's education seems to be less important for girls' educational outcomes, or put alternatively, that it reduces the deficit of boys. In other words, the positive effect of having a matric-educated mother is smaller for girls than it is for boys. Not only does this seem counterintuitive, but results which seem to suggest precisely the opposite have generally been found in the literature.

None of the coefficients on the interaction terms between paternal education and the female dummy are significant, suggesting that fathers' education impacts the grade attainment of boys and girls in much the same way. This is also somewhat surprising given findings of Glick

& Sahn (2000) of the importance of paternal education for boys' schooling outcomes especially. One potential explanation for this insignificant coefficient may lie in the large proportion of respondents who live in father-absent households. If father co-residence is a good measure of father involvement, the fact that so many respondents do not reside with their fathers may mean a similarly large proportion do not interact with their fathers enough for their fathers' education to impact on their own years of schooling completed. To investigate how the effect of father's education differs for absent versus co-resident parents, Model 7 includes interaction effects between parental co-residence and parental education.

Some interesting results emerge (Model 7's results are reported in column 4 of Table 3). Firstly, even though mother's co-residence in female-headed households is insignificant in both the fifth and sixth models, it is large and significant in the seventh model, suggesting respondents who live in female-headed households where their mothers are present have obtained, on average, 0.392 more education than their counterparts from mother-absent households. In other words, mother co-residence becomes significant only once one includes interaction terms between mother co-residence and mother's education. This may suggest that the effect of a mother's residence in the household on child educational outcomes differs based on the mother's level of education. Specifically, as is evidenced by the coefficients on the interaction terms between maternal education and mother co-residence, a mother's residence in the household appears to work in the opposite direction of the positive effect of her education. Respondents who live in households with their mothers who have completed primary school have themselves completed 0.395 *less* education relative to respondents who reside in female-headed mother-absent households. Similarly, respondents who reside in mother-present female-headed households where their mothers has obtained some secondary schooling have completed 0.388 less education than respondents whose mothers have completed some secondary education and are *not* resident in their households. The apparent negative effect of a mother's co-residence when interacted with her education is even larger for mothers who have completed matric or more: respondents in female-headed household whose mothers have matric or more and are resident in their household have attained 0.417 less education, on average, than those whose mothers have matric or more but are not co-resident. These constitute some very puzzling results. The consensus in the literature is that a mother's residence in the household is associated with positive educational outcomes for her children. Heard (2007: 437), for example, explains that mothers are more likely than other caregivers to participate in time-intensive activities with their children, such as school

interaction, monitoring of educational progress, and cognitive activities. The presence of more educated mothers, especially, has been shown to be positively associated with a number of child outcomes, including performance in education. Heckman (2008: 317) reports results from time diary studies that college-educated mothers devote more time to child-rearing than less-educated mothers, especially in child enrichment activities.

These findings in the literature suggest it is not just having an educated mother, but having an educated mother who is resident in your household especially, that has a positive effect on one's educational outcomes. The results presented here, on the other hand, suggest not only that such positive effects are not observable, but that a mother's residence in the household actually has a negative effect on grade attainment when interacted with mother's education. This suggests that the mechanism which causes a mother's education to impact positively on respondents' grade attainment is opposed by the mother's residence in the household. It must be kept in mind that this is only the case for female-headed households, where the father is absent by implication. In other words, living with a single mother who is educated seems to decrease the positive effect of her education for respondents living in female-headed households. It is not therefore the presence of an educated mother in the household per se which seems to decrease the positive effect of her education on her child's outcomes, but rather the presence of an educated mother in a female-headed household specifically, i.e. one where the father is absent, which can be said to decrease this positive effect.

Similar negative effects are not observable for paternal education. The only interaction term that has a statistically coefficient is that between a father who has completed some secondary education and his residence in the household, and this effect is in the direction one would expect. The coefficient suggests respondents who have fathers who have completed some secondary education and are resident in their households have attained 0.289 more education than their counterparts whose fathers have completed some secondary schooling but are absent from their households.

Although perhaps not as counterintuitive as the coefficients on the interaction terms between maternal co-residence and maternal education, the coefficients on the interaction terms between paternal co-residence and education are nonetheless somewhat surprising. The fact that all but one are statistically insignificant suggests that on average, the positive effect of a father's education on their child's grade attainment is not enhanced by having both parents resident in the household. As is the case with the counterintuitive results on the interaction

terms with mother co-residence and maternal education, this suggests that whatever the mechanism which causes a father's education to impact positively on his child's schooling outcomes, it does not require the parents' residence in the household to work.

One potential explanation for the negative association observed between a mother's co-residence and her education may be that more educated mothers are absent from their children's households due to participation in the labour market. The fact that labour migration is common in South Africa, as discussed in the literature review, certainly makes the idea that labour market participation would be associated with being absent from the household somewhat plausible. To investigate whether such effects may be present, the seventh model includes controls for whether parents have ever worked. The results are presented in column 2 of Table 4.

The first thing to note about Model 8 is the decrease in sample size from the previous model specifications (5,764 observations in the previous model compared with 3,695 in this one). This is due to the parental occupation variable having many missing values. The resulting sample is, like the previous samples, biased, this time only including respondents for whom we have information about parental education as well as occupation. Once again, it is important that this be kept in mind in the discussion that follows.

Neither the coefficient on having a mother who has never worked, nor having a father who has never worked is significant. In other words, all else being equal, respondents whose parents have never worked are not statistically significantly different in terms of grade attainment from those whose parents have been employed at some point in time. It must be noted that this particular measure of parents' labour market participation is extremely crude, given that a parent might have been unemployed at the time of the survey, or indeed for most of their adult life, and would still not be recorded as never having worked if they had held a job for any period of time at some stage in their lives. Nonetheless, the fact that about a third of respondents have at least one parent who has never worked indicates this variable does distinguish between respondents in at least a somewhat meaningful way.

Table 4: OLS regression results of parent occupation model specifications

	Years of education (8)	Years of education (9)	Years of education (10)
Age	0.684 (21.36)**	0.686 (21.43)**	0.684 (21.33)**
Female	0.595 (12.19)**	0.502 (7.81)**	0.595 (12.18)**
White	-0.167 (0.81)	-0.163 (0.79)	-0.168 (0.82)
Coloured	-0.063 (0.83)	-0.064 (0.84)	-0.058 (0.76)
Indian	0.097 (0.34)	0.081 (0.28)	0.105 (0.36)
Income (thousands of Rand)	0.078 (2.88)**	0.077 (2.84)**	0.080 (2.92)**
Income^2 (thousands of Rand)	-0.000 (1.81)	-0.000 (1.76)	-0.000 (1.84)
Mother resident in household	0.032 (0.56)	0.031 (0.53)	-0.022 (0.31)
Mother completed primary schooling	0.179 (2.12)*	0.182 (2.17)*	0.178 (2.12)*
Mother has completed some secondary schooling	0.479 (5.72)**	0.483 (5.77)**	0.480 (5.73)**
Mother completed matric or more	0.875 (9.36)**	0.882 (9.44)**	0.875 (9.36)**
Mother has never worked	-0.037	-0.189	-0.116
Constant	-3.517 (6.62)**	-3.504 (6.59)**	-3.478 (6.53)**
R^2	0.20	0.20	0.20
N	3,695	3,695	3,695

* $p<0.05$; ** $p<0.01$

Table 4: OLS regression results of parent occupation model specifications (cont.)

	Years of education (8)	Years of education (9)	Years of education (10)
Father resident in household	0.131 (1.98)*	0.132 (1.99)*	0.143 (2.04)*
Father completed primary schooling	0.006 (0.08)	0.007 (0.10)	0.005 (0.07)
Father completed some secondary Schooling	0.231 (2.89)**	0.233 (2.92)**	0.230 (2.88)**
Father completed matric or more	0.393 (5.31)**	0.395 (5.35)**	0.392 (5.30)**
Father has never worked	-0.088 (1.52)	-0.088 (1.11)	-0.069 (1.12)
Female*[Mother has never worked]		0.306 (2.84)**	
Female*[Father has never worked]		-0.013 (0.12)	
[Mother has never worked]*[Mother resident in household]			0.159 (1.41)
[Father has never worked]*[Father resident in household]			-0.118 (0.63)
Constant	-3.517 (6.62)**	-3.504 (6.59)**	-3.478 (6.53)**
R^2	0.20	0.20	0.20
N	3,695	3,695	3,695

* $p<0.05$; ** $p<0.01$

That the coefficients on these variables are not significant might indicate that the effect of having a chronically unemployed parent differs either by the gender of the respondent or whether the parent is co-resident in the household or not. Certainly, the fact that these factors seem to influence the effects of parents' education on grade attainment suggests that they might be important here, too. To test this, two additional models were run which interact the variable indicating whether a respondent's mother or father has ever worked with the gender of the respondent and the co-resident of the parent respectively. The results of these regressions are reported in columns 3 and 4 of Table 4 (Models 9 and 10).

Interestingly, the coefficient on having a mother who has never worked now becomes significant at the 90% level, and indicates that respondents whose mothers have never worked

have attained 0.189 less education on average than those whose mothers have worked at some point in time. The coefficient of the interaction term between the dummy indicating whether the mother has ever worked and the female dummy is positive and significant, and suggests girls whose mothers have never worked have attained 0.306 more education than boys whose mothers have never worked. It therefore seems that being female opposes the negative effect of having a mother who has never worked, so much so that the net effect of being female and having a mother who has never worked is positive. This once again constitutes a somewhat unexpected and counterintuitive result. The literature suggests that if there is any gender effect on the negative influence of a mother who has never worked, it would be that girls are affected even more than boys. The fact that the opposite seems to occur here is therefore puzzling. The coefficient on the dummy indicating whether a respondent's father has ever worked remains insignificant in the ninth model. This might suggest having a father who has never worked does not have a statistically significant impact on grade attainment, even when allowing for potential effects to differ by the gender of the respondent. Alternatively, this may suggest any potential effect of having a chronically unemployed father may already be captured by other variables in the model. Given the high correlation between unemployment and educational attainment in South Africa, it is likely that unemployed fathers in this sample are also those with few years of education. The explanatory variables relating to paternal education may therefore capture potential 'effects' of unemployment. While one can only speculate about this with the results at hand, what we can say from the present analysis is that having a chronically unemployed father has no additional negative effect, which is not captured by the influence of paternal education, on respondents' grade attainment, on average.

The tenth model includes interaction terms between whether a parent has ever worked and their co-residence in the respondent's household. This was done to determine whether having a parent who has never worked is significantly associated with grade attainment once one allows for this effect to differ depending on whether the parent in question is resident in the respondent's household. The results are reported in column 4 of Table 4. Interestingly, the coefficient on having a mother who has never worked loses significance when including such interaction effects. The coefficient on having a father who has never worked remains insignificant. It is also interesting that neither the interaction term between having a mother who has never worked and mother co-residence nor the interaction term between having a father who has never worked and father co-residence is significant. This suggests the effect –

or lack thereof – of having a parent who has never worked does not change depending on whether the parent is co-resident in the household. In other words, respondents whose mothers have never worked and are resident in their households are not statistically significantly different for those whose mothers have never worked and are absent from their households. The same is true for fathers.

5. Discussion

5.1. Household composition

The presence of both parents in the household was positively associated with the grade attainment of respondents in eight out of nine models, suggesting that the residence of both biological parents in the household is important for children's educational outcomes. Mother's co-residence, on the other hand, was found to be positively associated with grade attainment in only one out of the nine models in which it was controlled for. A likely explanation for this is that the effect of mother's co-residence was only tested in female-headed households. Since such households are typically associated with a number of vulnerabilities which impact negatively on children's schooling outcomes, it is likely that any positive effect of a mother's residence in the household is masked by the negative effect of female headship and the absence of the father.

5.2. Parent characteristics

The results show a clear association between the education of parents and grade attainment. As discussed in the literature review, one potential explanation for this is that genetics, both in the form of IQ as well as non-cognitive skills, play an important role in the transmission of educational attainment from parents to children.

Another possible mechanism discussed in the literature review is that more educated parents, especially mothers, raise the IQ and non-cognitive skills of their children by spending more time in child enrichment activities than their less educated counterparts. Interaction terms between parental education and co-residence were introduced as a way of testing this hypothesis. If the mechanism whereby parents' education influences their children's education is the type of activities they engage their children in, this effect should be stronger for co-resident parents since they spend more time with their children than non-co-resident parents. Surprisingly, the coefficients on interaction terms between mother co-residence and mother's education are negative for all levels of maternal education, suggesting that a mother's residence in the household counters the positive effect of a mother's education on

her child's grade attainment. Once again, this effect was only tested for maternal education in father-absent households, which are also female-headed households. It is therefore likely that the negative effects of female-headship also mask the positive effects of mother co-residence in these models.

By contrast, interaction terms between father co-residence and father's education are insignificant for all levels of father's education, barring matric or more. This suggests there is no statistically significant difference between the effect of an absent versus co-resident father's education on their child's years of schooling attained for fathers who have not completed matric. In other words, a father's education seems to have the same association with his child's education, regardless of whether he is resident in the child's household or not. This is an especially interesting result, since it seems to suggest the mechanism whereby a father's education impacts on his children does not require his presence in the household to be effective. There are a number of data concerns regarding paternal education in the analysis at hand, however, which could possibly compromise the accuracy with which this association is estimated. As such, it would be premature to draw a definitive conclusion about the exact mechanisms whereby a father's education impacts on his children's schooling outcomes from the results presented above. What can be stated with certainty is that there is a clear relationship between paternal education and grade attainment for the sample at hand.

These limitations notwithstanding, the analysis seems to provide some support for the idea that parental education impacts on children's schooling outcomes through a 'pure' genetic effect. This would explain why parent co-residence does not seem to matter for the transmission of parental education to their children, since parents pass on their characteristics to their children through their genes, and their residence in the child's household is not necessary for the expression of these genes. The fact that the positive effect of a parent's education does not seem to be contingent on their residing in the same household as their child could mean their level of education simply reflects their ability to succeed in education, which they pass on to their child genetically. Their residence in the household would thus not be necessary for this transmission mechanism to be effective. This would explain why a parent does not need to be present in the household for their education to impact positively on their child: it is not in fact their education which directly influences the outcomes their child achieves. Rather, their education is simply correlated with their ability – cognitive and/or non-cognitive – which is the real mechanism whereby they influences their child.

This explanation assumes parent co-residence is a reliable measure of parental involvement. As explained earlier, this need not be the case. Madhavan et al (2014) provide an explanation for why parental involvement may not be accurately captured by parent co-residence which is especially applicable in the South African context, namely labour migration. If some of the parents of the many children who are recorded in the NIDS data as not residing in their children's households are in fact temporary labour migrants, this would mean that they do return to the household periodically. Such mothers and fathers would be quite different in terms of the influence they have on their children from parents who are completely absent from the household in the traditional sense.

Perhaps even more noteworthy is that the labour migration story may find some support in the negative coefficients observed between maternal education and mother co-residence (Model 7). These coefficients suggest mother co-residence is in fact negatively correlated with grade attainment. If the reason for mothers' absence is that they are migrant labourers, this absence might be positively correlated with children's grade attainment. Such positive effects could occur through a number of channels. Firstly, mothers who are absent due to work are likely to contribute more to household income than mothers who are unemployed, for example. Secondly, mothers who are migrant labourers may influence their children's grade attainment positively through the aspirations they encourage in their children. It is at least conceivable that having a mother who is part of the formal labour market increases children's aspirations for their own labour market outcomes and thereby motivates them to succeed in school. What is especially relevant here is that the combined effect of maternal education and mother co-residence was only tested for children residing in female-headed households. Since women are by definition the primary income earners in these households, it is more likely that mothers in female-headed households participate in the labour market than is the case for mothers residing in male-headed households. The genetics mechanism could also explain this observed association: mothers who succeed in the labour market may possess cognitive as well as non-cognitive abilities which are favoured in the labour market. They may pass this on to their children genetically, who in turn perform better in school due to these abilities.

Such a possible labour market effect was partially tested for in the eighth, ninth and tenth models of the analysis by controlling for whether a parent has ever worked. The coefficients on these variables proved to be insignificant for the sample as a whole. Part of the reason for these insignificant coefficients may be that a variable indicating whether a mother or father has ever worked is a very crude measure of their labour market participation. Unfortunately,

given the difficulty of interpreting coefficients on variables with more information about parents' occupational status in the NIDS data, other measures of parents' labour market status could not be included here.

While the labour migration story may provide a relatively intuitive explanation for some very counterintuitive results, it is not possible to establish whether these effects are in fact driving the observed results given the limited information on parents' occupation in the data at hand.

The conclusions that can be drawn from the present analysis regarding the association between parental involvement and educational outcomes is therefore limited. We can only say with certainty that only the presence of both parents in the household is positively associated with the years of schooling a child attains, and that the education of both parents has a positive association with the child's years of schooling. This association is separate from the positive effect of residing in a household where both parents are present. Furthermore, maternal education appears to be twice as important for children's schooling outcomes as paternal education. Lastly, having a mother or father who has never worked does not seem to be significantly associated with the grade attainment of children.

6. Limitations

A major limitation of the present study is the cross-sectional way in which the NIDS data is used. Although using the panel as a cross-section greatly simplifies the analysis, one loses a lot of nuance in the association between home background and educational outcomes when doing so. Perhaps the biggest concern in this regard is that home background is not static - it is likely that respondents grew up in households that look very different from those captured at the time of the survey. Parental co-residence is of specific concern here. The analysis at hand implicitly assumes if a respondent resides in a mother-absent household at the time of the survey, they have never lived with their mothers. This is clearly a very problematic assumption, and may pose an explanation for some of the counterintuitive results that emerge from the analysis. But this deficiency is one that is also shared by most other studies on this issue that use cross-section data. The construction of a panel dataset which contains information about the home background as well as educational outcomes of South African children is therefore a crucial next step for further research on the role of home background in influencing educational outcomes in South Africa.

Furthermore the regression analysis presented here cannot account for the endogeneity that invariably plagues analyses of this sort. Regression analysis can simply reveal whether an association exists between given variables. While establishing an association between home background and schooling outcomes in South Africa is an important first step in understanding the mechanisms which cause this association, no causal inferences can be made from the present analysis. Establishing causality behind the associations uncovered in this study is therefore an important avenue for future research.

7. Conclusion

This paper has investigated the relationship between home background and grade attainment for a sample of 15-17-year-old respondents in NIDS. The presence of both biological parents in the household was shown to be significantly associated with the highest grade attained by respondents. Furthermore, a clear association exists between respondents' grade attainment and the level of education their parents have completed. The association between maternal education and grade attainment is almost twice as large as that between paternal education and grade attainment. This remains true even when one controls for mother and father co-residence. There are a number of transmission mechanisms that may cause these associations. Unfortunately, the NIDS data contains insufficient information about parental involvement to untangle these different chains of causality. Given the importance of isolating these transmission mechanisms for effective policy design, it is crucial that future research explore the possible causal relationships behind the associations uncovered in this paper.

References

Anderson, K.G., 2005. Relatedness and investment in children in South Africa. *Human Nature*, 16(1), pp.1–31.

Anderson, K.G., Case, A. & Lam, D., 2001. Causes and Consequences of Schooling Outcomes in South Africa. *Social Dynamics*, 27(January 2015), pp.37–59.

Buis, ML, 2012. The composition of family background: The influence of the economic and cultural resources of both parents on the offspring's educational attainment in the Netherlands between 1939 and 1991. European Sociological Review 29(3), 593–602.

Bowles, S. & Gintis, H., 2002. The Inheritance of Inequality. *Journal of Economic 16(3), 3–30.*, 16, pp.3–30.

Bowles, S., Gintis, H. & Osborne Groves, M. eds., 2005. *Unequal Chances: Family Background and Economic Success*. New York: Russell Sage Foundation.

Glick, P. & Sahn, D.E. 2000. Schooling of boys and girls in a West African country: the effects of parental education, income and household structure. *Economics of Education Review.* 19(1): 63-87.

Handa, S. 1994. Gender, Headship and Intrahousehold Resource Allocation. *World Development.* 22(10): 1535-1547.

Heard, H.E. 2007. Fathers, Mothers, and Family Structure: Family Trajectories, Parent Gender, and Adolescent Schooling. *Journal of Marriage and Family.* 69(1): 435-450.

Heaton, T.B.., Amoateng, A.Y.. & Dufur, M.., 2014. Race differences in educational attainment of youth aged 7-18 in post-apartheid South Africa: The role of family structure, resources and school quality. *South African Review of Sociology*, 45(1):101–121.

Heckman, J.J., 2008. Role of income and family influence on child outcomes. *Annals of the New York Academy of Sciences*, 1136, pp.307–323.

Madhavan, S., Gross, M.C., Richter, L., Norris, S. & Hosegood, V. 2014. Mothers, Kin, and Father Involvement in Urban South Africa. *Maryland Population Research Centre Working Paper.*

Mboya, M. M. & Nesegani, R.I. 1999. Migrant Labour in South Africa: A Comparative Analysis of the Academic Achievement of Father-Present and Father-Absent Adolescents. *Adolesence.* 34(136): 763-767.

Noble, K.G., Houston, S.M., Kan, E. & Sowell, E.R. 2012. Neural correlates of socioeconomic status in the developing human brain. *Developmental Science.* 15(4): 516-527.

Posel, D., 2010. Who are the heads of household , what do they do , and is the concept of headship useful ? An analysis of headship in South Africa. *Development Southern Africa.*18(5): 37–41.

Spaull, N., 2012. Equity & Efficiency in South African Primary Schools A Preliminary Analysis of SACMEQ III South Africa University of Stellenbosch Department of Economics.

Townsend, N., Madhavan, S., Tollman, S., Garenne, M. & Kahn, K. 2002. Children's residence patterns and educational attainment in rural South Africa, 1997. *Population Studies.* 56(1): 215-225.

Van Der Berg, S., 2007. Apartheid's enduring legacy: Inequalities in education. *Journal of African Economies*, 16(October 2005), pp.849–880.

Van der Vleuten, L. 2016. Mind The Gap! The Influence of Family Systems on The Gender Education Gap in Developing Countries, 1950–2005. *Economic History of Developing Regions*. 31(1): 47-81.

Van Wyk, C. 2015. An overview of Education data in South Africa: an inventory approach. *Stellenbosch Economic Working Papers.*

Zimmerman, F.J., 2002. Cinderella Goes to School: The Effects of Child Fostering on School Enrollment in South Africa. *Journal of Human Resources*, 38(3), pp.557–590.

Zoch, A., 2015. Life chances and class: Estimating inequality of opportunity in South Africa for various life stages. *Development Southern Africa*, 32(May), pp.57–75.