HIV/AIDS and Educational Attainment Among South African Children and Adolescents

Research Report (ECON7053A)
submitted in partial fulfilment (50%) of the
Degree of Master of Commerce (CCA01 – Health Economics)
in the School of Economics and Finance,
University of the Witwatersrand

by

Simon Halvey
2029536
simon.t.halvey@gmail.com

Supervised by Professor Frikkie Booysen

June 2023

Word Count: 17,205

Abstract

This study examines the effects of HIV-infection, antiretroviral treatment exposure and viral load suppression on indicators of educational attainment among children and adolescents of school-going age in South Africa using data from a nationally representative survey and linked biomedical information. This study employs propensity score matching-based approaches to examine the effects of HIV-related indicators on educational attainment overall and within each gender. This study also uses decomposition analyses to examine the contribution of HIV-related indicators to gender gaps in educational attainment, and to identify factors contributing to differences in the impacts of HIV-related factors on educational attainment of females specifically. It is found that HIV-positive status has a significant impact on the likelihood of school attendance, absenteeism, and being old for one's grade. This is the case for all children and adolescents of school-going age. These effects are statistically significant for girls but not for boys. The study also finds that HIV-related factors contribute to gender gaps in school attendance, and that these gaps are observed in all tiers of the HIV treatment cascade. These results indicate that HIV-related factors impact disproportionately and negatively on educational attainment for girls. Further research on the gendered nature of the impact of HIV on education in South Africa is encouraged, as is the design of health and education policies that specifically target improved education outcomes among HIV-positive girls.

Table of Contents

List of Figures

List of Tables

Table of Acronyms

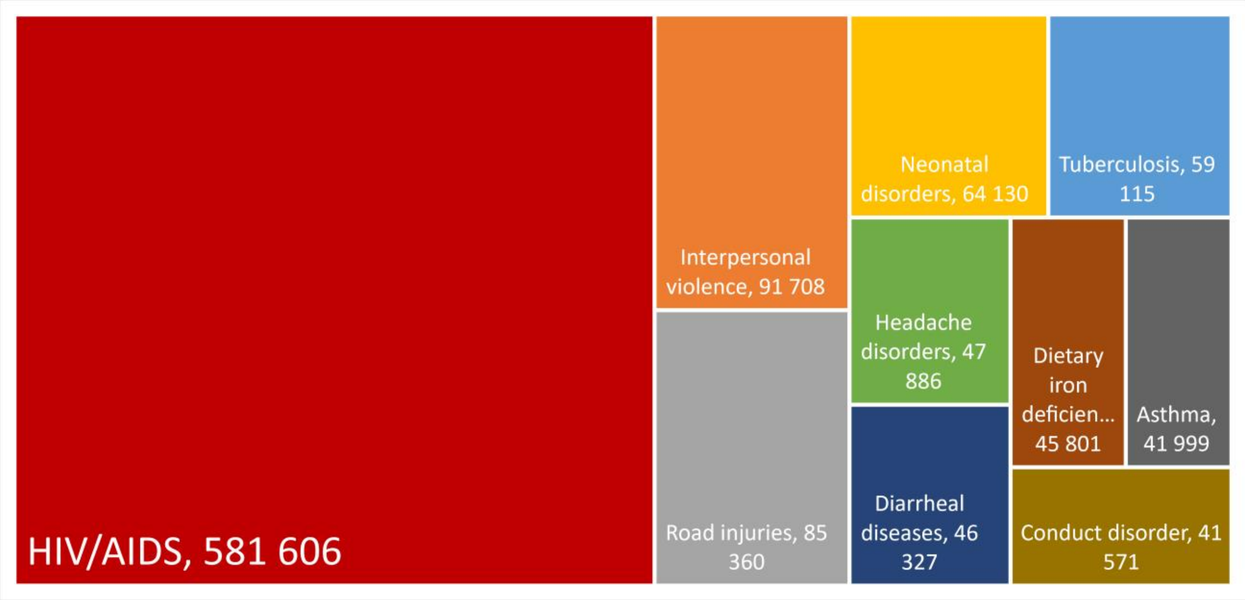
AIDS	acquired immunodeficiency syndrome
ART	antiretroviral treatment
DALY	disability-adjusted life year
HIV	human immunodeficiency virus
HSRC	Human Sciences Research Council
RCT	randomized controlled trial
RNA	ribonucleic acid
SABSSM	South African National HIV Prevalence, Incidence, Behaviour and Communication Survey
STI	sexually transmitted infections
UNAIDS	Joint United Nations Programme on HIV/AIDS

1 Introduction

1.1 Background

Despite the significant gains made in curbing the HIV/AIDS pandemic over the past two decades, the disease remains a significant public health concern in South Africa – HIV prevalence rates in South Africa are the fourth-highest globally, while the total population living with HIV in the country – an estimated 7.9 million individuals, representing nearly 20% of the global population living with HIV – is by far the highest globally. Indeed, most recent estimates from the Global Burden of Disease Study suggests that HIV/AIDS remains the single largest cause of premature mortality and morbidity in the country. Of chief concern is a trend of burgeoning HIV-related premature mortality and morbidity among children and adolescents of school-going age (aged 5-19 years) in South Africa. Indeed, as illustrated in Figure 1, below, HIV/AIDS represents the largest contributor to the accrual of disability-adjusted life years (DALYs) – a measure of overall disease burden – among children and adolescents of school-going age in South Africa, accounting for nearly a third of total disease burden within this age cohort in 2019 (IHME, 2021). Moreover, HIV/AIDS is often a contributing comorbidity to many of the other significant contributors to morbidity among this age-group, most notably tuberculosis, which is the fifth-largest contributor, suggesting that HIV/AIDS is an even more significant cause of disease burden among children and adolescents of school-going age in South Africa (IHME, 2021).

Figure 1: Total disability-adjusted life years (DALYs) among children and adolescents of school-going age in South Africa, by cause, 2019



Source: IHME (2021)

This substantial disease burden is associated with significant losses in human capital – losses which, among younger cohorts, are a direct function of the lower levels of educational attainment typically associated with HIV-positive status among children. Evidence in the literature (Zinyemba, Pavlova and Groot, 2020; Goldberg and Short, 2016; Guo, Li and Sherr, 2012) suggests several mechanisms through which HIV can impact on a child's educational attainment: (1) a child may miss school days due to HIV-related illness or treatment-related issues; (2) HIV-positive parents may not be able to facilitate their children's schooling due to HIV-related morbidity; or (3) the well-documented HIV-poverty nexus may contribute to socio-economic barriers to educational enrolment and attainment.

1.2 The 90-90-90 strategy

The first of these mechanisms suggests that increased severity of HIV-related morbidities is associated with increased levels of school absenteeism and associated reductions in educational attainment. In terms of HIV-related morbidities, specifically, the rate and severity of said morbidities is a direct function of the HIV viral load – the number of HIV copies present in a millimetre sample of blood (CDC, 2022). Such high HIV viral load is a chief determinant of the severity of morbidity in a patient. As a result, achievement of viral load suppression – or the reduction of HIV viral load to the point (under 200 copies per millimetre) that the individual's immune system remains uncompromised and the virus is no longer transmissible

– through strict adherence to an antiretroviral treatment (ART) regime – remains the foundation of HIV treatment and prevention interventions globally (CDC, 2022).

To this end, in 2014, the Joint United Nations Programme on HIV/AIDS (UNAIDS) implemented the 90-90-90 targets: namely, that, by 2020, 90% of people living with HIV know their HIV status (the first 90); that 90% of all people with diagnosed HIV infection receive sustained antiretroviral therapy (the second 90); and that 90% of all patients receiving antiretroviral treatment are virally suppressed (the third 90). Achievement of this cascade of targets by 2020 would result in at least 73% of people living with HIV achieving viral load suppression, with modelling from UNAIDS suggesting this would sufficiently reduce the number of new infections so as to achieve HIV/AIDS epidemic control by 2030 (UNAIDS, 2014). Granted the ubiquitous nature of the 90-90-90 targets in HIV treatment and prevention campaigns globally, and granted the fact that these targets specifically seek to achieve viral load suppression which manifests as reduced rates and severity of HIV-related morbidities in individual patients, thereby impacting on educational attainment and human capital development, the three targets of the 90-90-90 cascade – knowledge of HIV status, exposure to ART, and viral load suppression – therefore represent a logical set of independent variables to model the impact of HIV on educational attainment in South African children and adolescents of school-going age.

Thus, there is ample theoretical basis for interrogation of the impact of antiretroviral treatment exposure and adherence, and associated viral load suppression status, among HIV-positive children of school-going age on educational attainment. Indeed, while there is an established literature on the human capital loss brought about by HIV in South Africa specifically, and in sub-Saharan Africa more generally, there is a dearth of literature on the impact of specific HIV-related indicators – namely antiretroviral treatment (ART) access, ART exposure, and viral load suppression – on educational attainment rates specifically. Moreover, there is little evidence of how gender may contribute to the manifestation of these effects, and, specifically, whether HIV has a larger impact on educational attainment rates among girls than it does among boys.

Given the aforementioned gap in the literature, this study, employing data with biomedical information on HIV for children and adolescents of school-going age in South Africa from the

5th South African National HIV Prevalence, Incidence, Behaviour and Communication Survey (SABSSM V), conducted in 2017/8 by the Human Sciences Research Council, uses propensity score matching-based analyses to examine the impact of HIV status, exposure to ART, and HIV viral load suppression on educational attainment among HIV children in South Africa and to examine how these impacts differ across genders. Further to this, the study also uses Oaxaca-Blinder decomposition analysis to interrogate the extent to which these HIV indicators help explain gender gaps in educational attainment, and to assess the different contributions to HIV-related educational attainment gaps within each gender.

2 Theoretical Framework

The relationship between health status and education is well established in the social science literature. Indeed, so prevalent and strong is this relationship, that Kolata (2007) notes there is near universal agreement among researchers that education, and educational attainment specifically, is the single most consistent and significant correlate of good health globally. Attempts to interrogate and explain this long-identified relationship were formalized by Michael Grossman (1976) who sought – based on his seminal model of health demand, which posited health as a durable capital good – to develop a methodological framework under which various explanations of the correlation between health and schooling could be examined (Grossman, 1976). Broadly, Grossman acknowledges three chief explanations for the established positive correlation between health and education: one which posits that the causal relationship runs from increases in education to increases in health status; a second which posits that the causal relationship runs in the opposite direction, from improved health status to increases in educational attainment; and a third which holds that there is no causality between the two variables, but rather that differences in one or more external variables affect health and education in the same direction (Grossman, 1976). Under the second of these assumptions, the direct effects of poor health on education are patently obvious – a child in poor health is almost certain to be absent from school more frequently than a child in good health, and is also likely to learn less in the days when they are at school as a function of this absenteeism and established morbidity-related obstacles to learning.

However, Grossman (1976) notes a further potential effect of poor health on educational attainment, based on a model of demand for schooling developed in the same vein as his health demand model. Here, Grossman notes that demand for schooling – which, like health, is a human capital investment – is a positive function of the efficiency with which a child is able to transform education into a stock of knowledge. This efficiency in the production of human capital from education is a direct function of a child's mental ability and, of key importance to this study, physical health status (Grossman 1976; Lillard, 1973). The poorer a child's physical health, and by virtue thereof, the lower a child's efficiency in the production of human capital from investment in education, the lower the demand for investment in their education will be. Moreover, since education is an investment in human capital, the investment decision is likely to be determined by the number of periods over which returns on the investment can be collected. In the case of education, this number of periods is directly determined by the number of years that a child is expected to work and earn a wage, using their knowledge stock to generate income; these number of years are, in turn, a function of the child's future life expectancy and, by virtue thereof, their health status – the poorer a child's health status, the lower the expected return on an investment in their education, and the lower the demand for such an investment (Grossman, 1976). Thus, in addition to the effect of a child's health status on their educational attainment, as a function of enforced absenteeism and poorer learning while at school, poor health status may have a negative impact on the *demand* for education, thus further impacting on their potential educational attainment. It is both of these potential effects which form the theoretical basis for the empirical hypotheses investigated in this study.

3 Literature Review

A review of the literature suggests that several studies globally have interrogated the direct effects of HIV on children's educational attainment. A key contribution to this literature is the systematic review of such studies by Zinyemba, Pavlova and Groot (2020), which identifies 62 studies – covering quantitative, mixed methods and qualitative approaches – interrogating the impact of HIV on school attendance. Specifically, this systematic review distinguishes three main effects of HIV on educational attainment of children, namely, the effects on sick children and orphans, the effects on gender-gaps, and the effects on the intergenerational transmission of education. The former two relationships are of chief importance to this research.

In terms of the first effect – of HIV on educational outcomes of children – the authors identify 28 quantitative studies which interrogate this relationship specifically, 20 of which found that

HIV-affected children – that is children infected with HIV, children with HIV-positive parents, or children who have been orphaned due to HIV – attained less education than HIV-unaffected children (Zinyemba, Pavlova and Groot, 2020). A review of the individual studies that comprised this systematic review reveals that these lower levels of educational attainment among HIV-affected children were the manifestation of several direct health effects on children: namely higher rates of illness-related school absenteeism (Anabwani *et al.,* 2017; Mayes *et al*., 1996); more frequent incidence of school dropout among HIV-affected children (Parchure *et al*, 2016; Bele *et al*., 2011); and poorer educational performance (Ellis, 2004) due to illness.

Further systematic reviews have also contributed to the literature relevant to this discourse. A systematic review by Goldberg and Short (2016) summarizes research on children living with HIV-infected and AIDS-ill adults in sub-Saharan Africa, and on the impact thereof on child well-being specifically. Of the 45 empirical studies included in the systematic review, nine identify the effect of residing with an HIV-infected adult on childhood educational outcomes specifically. Across these studies, it was found that co-residing with HIV positive adults or parents is associated with lower school attendance – with a more significant impact at secondary school level (Floyd *et al*., 2009), higher rates of drop-out (Orkin *et al.*, 2014), non-enrolment (Cluver *et al*., 2013), and deficits in grade progression (Akbulut-Yuksel & Turan, 2013). These findings speak to one of the key mechanisms through which HIV infection can contribute to lower educational attainment, namely, that HIV-positive parents may not be able to facilitate their children's schooling due to HIV-related morbidity. Furthermore, several studies found that children of school-going age may miss school, drop out of school, or not be in the correct grade due to the need to provide care to parents infected with HIV (Pufall *et al*., 2014, Cluver *et al*., 2012). Once again, it was found that improved coverage of antiretroviral treatment could mitigate these impacts. Zivin *et al*. (2009) found that providing HIV-positive parents of Kenyan children with antiretroviral drugs led to a significant increase in weekly hours of schooling, while Delva *et al*. (2009) found that the positive health impacts of increased ART coverage among HIV-infected parents led to increased labour market participation, and avoidance of delays to their children's education.

A systematic review by Guo, Li and Sherr (2012), meanwhile, identifies and reviews 23 studies in the global literature on the effects of HIV/AIDS on children's schooling. This review reveals significant educational disadvantages among children affected by HIV (those children being orphaned or made vulnerable by HIV/AIDS), relative to HIV-unaffected children. These educational disadvantages are manifested across a variety of educational outcomes, including school enrolment and attendance (Case & Ardington, 2006; Yang *et al*., 2006); inappropriate grade level for age (Kasirye & Hisali, 2010; Ainsworth *et al*., 2005); and poor school performance (Tu *et al*., 2009). Further studies have interrogated potential protective effects of education on HIV infection – that is the potential for schooling and the status of being in school to protect against the likelihood of HIV infection – with a number identifying such potential protective effects (Santelli *et al*., 2013; *Pettifor et* al., 2012; Bärnighausen *et al*., 2007).

Within this broader literature, however, there is a significant lack of studies that specifically examine the effects of HIV serostatus among children on their educational attainment, largely due to a dearth of datasets that include sufficient biomedical information about children's HIV serostatus (Zinyemba, Pavlova and Groot, 2021). In one study which does examine such effects, Parchure *et al*. (2016) interrogate the educational outcomes of family-based HIV-infected children from Maharashtra, India, using data from a survey of 510 households in the region. The study found that HIV-infected children of school-going age were seven times more likely to be out of school than their HIV-negative counterparts, and were also three times more likely to be in an inappropriate grade for their age (Parchure *et al*., 2016). Another study by Anabwani *et al*. (2016) examines the schooling outcomes of HIV-infected children in Botswana, using data from a cross-sectional survey among HIV-infected and HIV-uninfected children aged 6-8 years. The authors find higher levels of school absenteeism among HIV-infected children, than their HIV-uninfected counterparts, while inappropriate grade level for age was three times more prevalent among HIV-infected children (Anabwani, *et al*., 2016). While this literature does present an initial exposition of the effect of HIV on children's educational attainment, what is apparent in the literature is a scarcity of studies that examine the effect of childhood HIV infection, specifically, due to a lack of datasets with appropriate information on children's HIV serostatus. The availability of such data in the SABSSM dataset (which is discussed in further detail below), allows for a fairly significant contribution to this literature.

In terms of the second main effect of HIV on educational attainment of children identified by Zinyemba, Pavlova and Groot (2020) – the effects on gender-gaps – the authors identify 10 papers in the literature which interrogate such effects, specifically, six of which find negative effects of HIV on gender differences in educational attainment (Zinyemba, Pavlova and Groot, 2020). The direction and magnitude of these effects on gender gaps, however, do differ across these studies: while some studies found that the impact of HIV of educational attainment was pro-boy, realized as higher rates of school attendance (Delva *et al*., 2009) or better school performance (Harrison *et al*., 2017) among HIV-affected boys than HIV-affected girls, other studies found that the effect was pro-girl in nature, with higher rates of grade progression (Orkin *et al*., 2014), higher likelihood of appropriate grade level for age (Pufall *et al.*, 2014), and higher levels of educational functioning (Hensels *et al*., 2016) among HIV-infected girls than HIV-infected boys. What is abundantly clear from this literature review, however, is that the extant literature on the direct impact of childhood HIV-infection on children's educational attainment, and on the effects thereof on gender gaps in educational attainment specifically, at present is rather limited.

4 Data

4.1 SABSSM Survey

This study uses data from the South African National HIV Prevalence, Incidence, Behaviour and Communication Survey (SABSSM) conducted by the HSRC in 2017, employing nationally and provincially representative samples. This cross-sectional, population-based household survey was conducted using a multistage stratified random cluster sampling approach. This survey included people of all ages living in South Africa. All members in the selected households were invited to participate. Data was collected from respondents using a set of four questionnaires – a household (or visiting point) questionnaire, and age-cohort specific questionnaires for parents and guardians of children aged 0-11 years, children aged 12-14 years, and people aged 15 years and older. Granted the sensitive nature of the data being collected, informed consent was sought from all respondents. To link survey data with health status indicators of interest, questionnaires were linked with anonymous testing procedures, with testing conducted at accredited laboratories.

The survey targeted an initial 15,000 visiting points, of which 11,776 were valid visiting points approached for participation, from which 9,656 agreed to participate in the survey. From these 9,656 visiting points, 39,132 individuals were eligible for interview and to provide a blood sample. Among these eligible individuals, 36,609 (93.6%) agreed to be interviewed, with the 6.4% non-response attributable to individuals being absent from the household at the time of the interview (2.4%), or refusing to be interviewed altogether (4.1%). Blood specimens were provided by 23,923 (61.1%) of eligible individuals and anonymously linked to completed questionnaires. A total of 12,686 individuals who completed the questionnaire refused to provide a blood sample, accounting for 32.4% of the eligible study population.

The resultant SABSSM dataset contains provincially and nationally representative data on a variety of HIV-related indicators. Of specific interest to this study, these include data from the blood specimens on HIV serostatus, exposure to antiretroviral treatment, viral load suppression, and HIV drug resistance. These data are linked with data from the various questionnaires covering demographics; orphan status; education; HIV knowledge, attitudes beliefs and values; sexual history; delivery and care details; male circumcision; HIV testing and risk perception; drug and alcohol use; health status; mental health; and intimate partner violence. This dataset, therefore, provides a large, robust and representative sample, using observational data, through which the causal effect of interest – namely of relevant HIV serostatus indicators on educational attainment – can be duly interrogated using appropriate econometric methods.

4.2 Causal Inference with Observational Data

While the SABSSM survey, as described above, clearly provides a large and variable-rich dataset, based on nationally and provincially representative samples, the fact remains that the data therein is observational in nature, which entails an important problem in the estimation of such treatment effects. Indeed, it is widely acknowledged that randomized controlled trials (RCTs) – an intervention study, as opposed to an observational study, in which the treatment of interest is randomly assigned to one of two groups, and the response thereto measured – represent the gold standard in research design for causal inference, by virtue of the elimination of various forms of bias that are inherent observational studies (Pan and Bai, 2018). However, the randomized controlled trial approach is not feasible in this instance, granted the clear ethical barriers to the systematic denial of life-saving antiretroviral treatment from one cohort of the population. In lieu of a feasible RCT-based research design, observational studies are often used to draw inferences from a sample to a population on the possible effect of a treatment on a treatment group versus a control group, where assignment of the independent treatment variable is beyond the control of the researcher. Naturally, since assignment of subjects to the treatment and control groups is not random in observational studies, the estimation of treatment effects under such studies are at increased risk of bias, due to the likelihood of systematic non-random selection of individuals into treatment or control groups, and, similarly, due to the potential existence of confounding factors (Nichols, 2007; Becker and Ichino, 2002).

A prerequisite for the robust estimation of treatment effects in observational studies, therefore, is the use of quasi-experimental research designs, which employ approaches to causal inference that, by their design, serve to reduce the bias inherent in observational studies. One such approach to reducing bias inherent in the estimation of treatment effects with observational datasets that has gained increasing traction in the health economics literature is propensity score matching. First espoused in the seminal paper by Rosenbaum and Rubin (1983), propensity score matching, as the name suggests essentially aims to mimic the properties of a randomized controlled trial by creating control and treatment groups that are comparable on all relevant observed covariates – measured as the conditional probability of assignment to a particular treatment given observed covariates – and then by estimating the average treatment effect as per typical experimental designs (Pan and Bai, 2018). As Caliendo and Kopeinig (2008) note, propensity score matching can be, and indeed has been, applied in a variety of fields of study in which causal treatment effects are being interrogated, including health economics. Furthermore, Caliendo and Kopeinig (2008) note that propensity score matching is applicable in all situations in which a study, akin to this one, has an identified treatment, an identifiable group of treated individuals, and an identifiable group of untreated individuals. Given the applicability of this approach to observational studies, to health economics specifically, and to datasets such as that being used for this study, propensity score matching clearly represents a feasible approach to interrogating the effect of HIV viral load suppression on educational attainment using the SABSSM dataset.

4.3 Examining Gender Gaps in Educational Attainment

Multivariate decomposition – first posited by Blinder (1973) and Oaxaca (1973) – is an approach which is widely used in economics to "quantify the contributions to group differences in average predictions from multivariate models" (Powers, Yoshioka and Yun, 2011). Specifically, this approach uses the outputs of a regression model to explain the difference in the means of a dependent variable between two groups by, as the name clearly suggests, decomposing such differences into two distinct parts in a counterfactual manner – one part that is due to *differences in characteristics*, and another part that is due to *effects of these characteristics*. This approach has been widely used to explore the determinants of gender gaps in educational attainment – Duryea *et al*. (2007) use decomposition analysis to examine the evolution of the educational gender gap in Latin America and the Caribbean over a forty-year period; Kingdon (2002) uses decomposition analysis to interrogate the reasons underpinning the education gender gap in Uttar Pradesh, India; and, most relevant to this study, Fotso, Banjo and Akinyemi (2018) use multivariate decomposition to interrogate the effect of the presence of HIV infection in the household on adolescents' educational attainment in South Africa. The use of multivariate decomposition in the estimation of gender gaps in educational attainment is, thus, well established, and the prerequisites for its implementation – a regression output with distinct outcomes by gender – are well supported by the SABSSM dataset.

5 Methodology

5.1 Identifying Relevant Variables in the SABSSM Dataset

Across the various analyses, the dependent variables of interest relate to school non-attendance (for children of school-going age), and overall educational attainment (for adolescents that are at an age where they ought to have completed schooling). As per the theoretical framework presented above, HIV infection could affect educational attainment through children being out of school altogether, children being in an inappropriate grade for their age, or through illness-enforced absenteeism. Thus, the dependent variables of interest, for all children and adolescents of school-going age (aged 5-19 years), are based on the following survey questions: "Are you currently attending school?"; "In the last month, have you missed school?"; "What grade are you attending this year?". The former two variables are already binary in nature, while the latter was combined with the child's age data to generate a binary variable to represent appropriate grade for age, with an age that is two years older than the typical age-for-grade level being deemed "inappropriate for grade level". For adolescents of possible matriculation age (17-19 years), meanwhile, the dependent variable of interest is based on the survey question: "What is the highest educational level that you obtained?", with a binary variable generated from this question to identify adolescents that had completed matric or higher. These binary variables, thus, represent the treatment variables employed in the propensity score matching analysis and as the determinants of gender gaps in educational attainment explored in the decomposition analysis. The chief independent variables include HIV status, ART status, viral load status, gender, age and relevant parental and household characteristics. The biomedical data are drawn from the biomedical testing results, while the remaining dependent variables are drawn from relevant survey responses.

5.2 Propensity Score Matching Analysis

As mentioned above, analysis of the causal effect of interest – namely of HIV infection and related indicators on educational attainment – is conducted using a propensity score matching-based approach. Inference about the causal effect of a treatment (e.g., HIV viral load suppression) on the outcome (e.g., school attendance) of an individual essentially involves speculation as to the potential outcome for said individual had they not received the specified treatment (e.g., what would the school attendance of an HIV-positive child without viral load suppression been had they been virally-suppressed?). Thus, under a binary treatment, such as that in this study, the treatment effect can be noted by the below equation:

Equation 1: Treatment effect for individual i

$$\tau_i = Y_i(1) - Y_i(0)$$

where τ is the treatment effect for individual i, $Y_i(1)$ is the outcome when individual i receives treatment (e.g., has not achieved HIV viral load suppression), and $Y_i(0)$ is the outcome when individual i receives no treatment (e.g., has achieved viral load suppression) (Caliendo and Kopeinig, 2008). Naturally, the above situation presents an impossibility, namely that it is impossible for an individual to simultaneously both receive and not receive a treatment – a child cannot be simultaneously HIV positive and negative, or be simultaneously viral load suppressed and non-suppressed. This encompasses the fundamental evaluation problem of the missing counterfactual – that for any individual only one of the two possible treatment scenarios and associated outcomes is observable – and necessitates the use of average treatment effects on the treated within a population, rather than an individual treatment effect, for interrogating the impact of a treatment of interest (Caliendo and Kopeinig, 2008). This average treatment effect on the treated is expressed thusly:

Equation 2: Average treatment effect on the treated

$$\tau_{ATT} \equiv E\{Y_1 - Y_0 \mid D = 1\}$$

This equation essentially reflects the difference between expected outcome values with and without treatment for those individuals within the population who are exposed to the treatment of interest. Naturally, this leaves the same problem as the individual treatment effect, as described above, namely, that the counterfactual outcome for those being treated – represented by $E[Y_0 \mid D = 1]$ in Equation 2 – is unobserved. This requires the development of a suitable substitute for this counterfactual outcome, such that the average treatment effect on the treated can be accurately estimated (Caliendo and Kopeinig, 2008). Propensity score matching seeks to substitute this missing counterfactual by identifying, within the population which did not receive the treatment, a group – hereafter known as the control group – which is as similar as possible to the group within the population which did receive the treatment – hereafter known as the treatment group – based on a selection of pre-treatment characteristics (Becker and Ichino, 2002). This approach essentially seeks to create separate control and treatment groups which are equivalent in their respective conditional probabilities of receiving treatment. This probability of receiving treatment is known as the propensity score and is represented by the below equation:

Equation 3: The propensity score

$$p(X) \equiv \mathrm{Pr}(\mathrm{D} = 1|\mathrm{X}) = \mathrm{E}(\mathrm{D}|\mathrm{X})$$

where X is the multidimensional vector of pre-treatment characteristics (Becker and Ichino, 2002). Substituting the propensity score from Equation 3 into the average treatment effect on the treated equation in Equation 2, allows for generation of an equation for the estimation of average treatment effects under a propensity score matching-based approach:

Equation 4: Average treatment effect on the treated under propensity score matching

$$\tau_{ATT}^{PSM} = E[E\{Y_1|D = 1, p(X)\} - E\{Y_0|D = 0, p(X)\}|D = 1]$$

The above equation, as a standalone, remains insufficient for the estimation of the average treatment effect on the treated, due to the continuous nature of the propensity score variable $p(X)$, which means that it would be impossible to find two observations with identical propensity scores (Becker and Ichino, 2002). In lieu of exact matching of propensity scores, a number of matching algorithms have been employed in the literature, so as to most accurately match observations in the treatment and control groups. The most commonly employed of these matching algorithms are: nearest-neighbour matching, the most straightforward matching estimator, in which a single individual from the control group is chosen as a match for a single treated individual by virtue of having the closest propensity score to said treated individual; kernel matching, a non-parametric matching approach, which uses the weighted average of all individuals in the control group to construct a counterfactual outcome; and stratification matching, in which the common support of the propensity score is partitioned into a set of strata and the mean difference in outcomes between treated and control observations are estimated within said strata. All three of these matching algorithms are applied in this study in order to interrogate the veracity and sensitivity of the results.

The propensity score matching approach, as outlined above, is implemented in this study using the *pscore* Stata package, as developed by Becker and Ichino (2002). This package is applied to the SABSSM dataset to generate an initial propensity score-based estimation of the average treatment effect on the treated for the whole population, which is used to interrogate the impact of HIV-related indicators (e.g., HIV viral load suppression) on indicators of educational attainment (e.g., school attendance) for the population as a whole; thereafter, separate PSM-based regressions for boys and girls are run in order to interrogate whether these impacts differ by gender, or specifically, whether HIV/ART/viral load-status affect educational outcomes differently for boys and girls. As a robustness check, these propensity score matching approaches are repeated in this study using two additional Stata packages: the *psmatch2* package, developed by Leuven and Sianesi (2003); and the *teffects* package, which is standard with Stata.

5.3 Decomposition Analysis

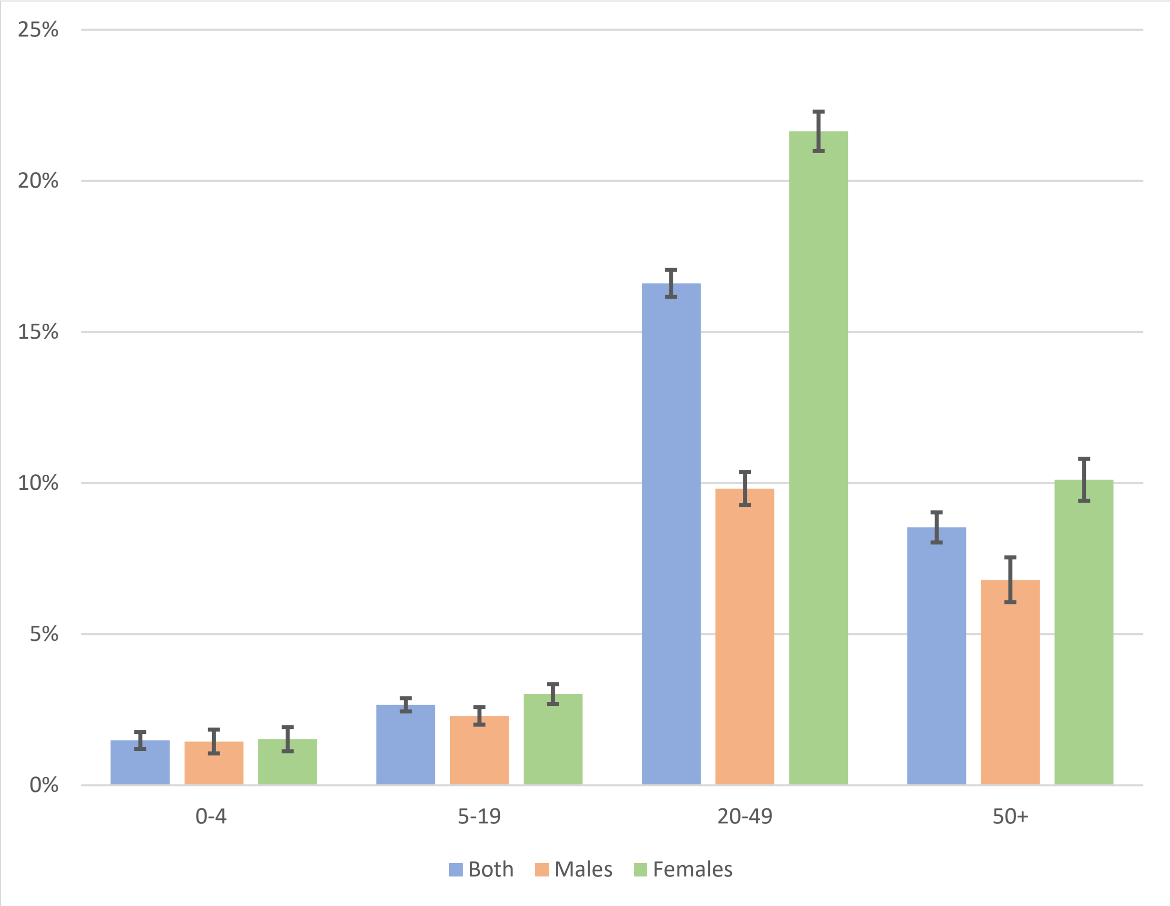
Also as mentioned above, analysis of the gender gaps in the educational attainment outcomes of interest and the extent to which these HIV indicators help explain gender gaps in educational attainment is conducted using the multivariate decomposition approach as espoused by Blinder (1973) and Oaxaca (1973). Jann (2008) has developed a Stata command, *oaxaca*, for such nonlinear multivariate decomposition, and it is this approach, and this Stata routine, which is employed in this section of the analysis. Specifically, the multivariate decomposition method developed by Jann (2008) is used to decompose the mean differences in the educational attainment variables as described above into differences in characteristics, which are described in the model as endowments, and differences in the effects of these characteristics, which are described in the model as coefficients. In order to fully analyse whether there are gender differences (HIV-positive boys vs. HIV-positive girls) in how the HIV indicators relevant to this study impact on gender gaps educational attainment, the following process is followed: an initial decomposition of the various measures of educational attainment by gender is conducted in the overall sample, with all three HIV treatment cascade indicators included as covariates in order to interrogate the potential contribution of these indicators to observed gender gaps in educational attainment; thereafter further decompositions of the various measures of educational attainment by gender are conducted within the various HIV status group sub-samples in order to interrogate whether, within the HIV-positive, ART exposed and virally suppressed sub-populations, respectively, there are any gender-based differences in the various educational attainment outcomes and the nature of these differences; thereafter, separate decompositions of educational attainment by HIV status, within each specific gender, are conducted in order to examine whether the impacts of these various HIV status indicators differ across gender and to interrogate the factors that contribute to these HIV status-based differences within each gender.

6 Results

6.1 Performance against the 90-90-90 targets

Figure 2, below, provides estimates from SABSSM of the proportion of people living with HIV in South Africa, disaggregated by gender and age cohort. As is evident in Figure 2, rates of HIV infection are highest among the 20-49 year-old age cohort (16.6%), and lowest among the 0-4 year-old age cohort (1.5%). HIV infection rates among children and adolescents of school-going age (5-19 years) are 2.7%. In terms of gender-based differences in rates of HIV infection it is notable that across all age cohorts, with the exception of the younger child cohort (0-4 years), rates of infection are significantly higher among females than among males – with 3.0% of females aged 5-19 years testing HIV positive, relative to 2.3% of males. Rates of infection among working age adults (20-49 years) are similarly higher among females (21.6%) than among males (9.8%); as are rates among older adults (50 years and older), where 10.1% of females are HIV-positive, relative to 6.8% of males.

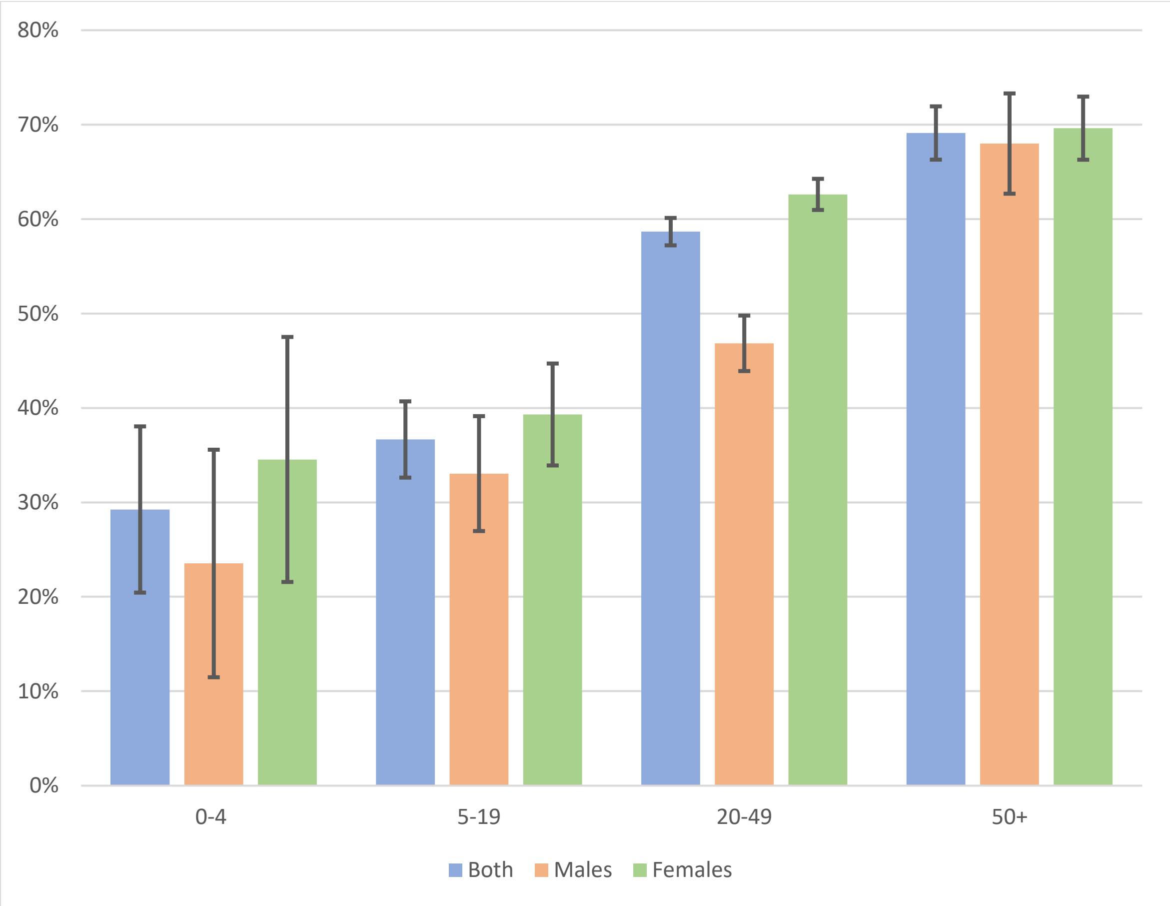
Figure 2: People living with HIV in South Africa, by gender and age cohort



Source: Own calculations from SABSSM 2017 dataset

Figure 3 presents an overview of the proportion of people living with HIV in South Africa who were on ART at the time of the SABSSM survey, disaggregated by gender and age cohort. As is evident in Figure 3, ART exposure is lowest among children and adolescents, with only 29.3% all children living with HIV aged 4 years and younger being exposed to ART. ART coverage rates among children and adolescents of school-going age living with HIV – those aged 5 to 19 years – are somewhat higher, at 36.7%, but still suggest that only slightly more than a third of all children and adolescents of school-going age in South Africa living with HIV are receiving ART. In terms of gender-based differences in rates ART exposure it is notable that across all age cohorts, mean rates of ART exposure are higher among females than among their male counterparts, although this gender difference is only statistically significant for working age adults (20-49 years). The low rates of ART coverage among the youth and adolescent cohorts in South Africa are an ominous harbinger for the rates of viral load suppression among the HIV-positive, and for the potential impact thereof on the health status of South African youth living with HIV, and, by extension, on their potential educational attainment.

Figure 3: ART exposure among people living with HIV in South Africa, by gender and age cohort

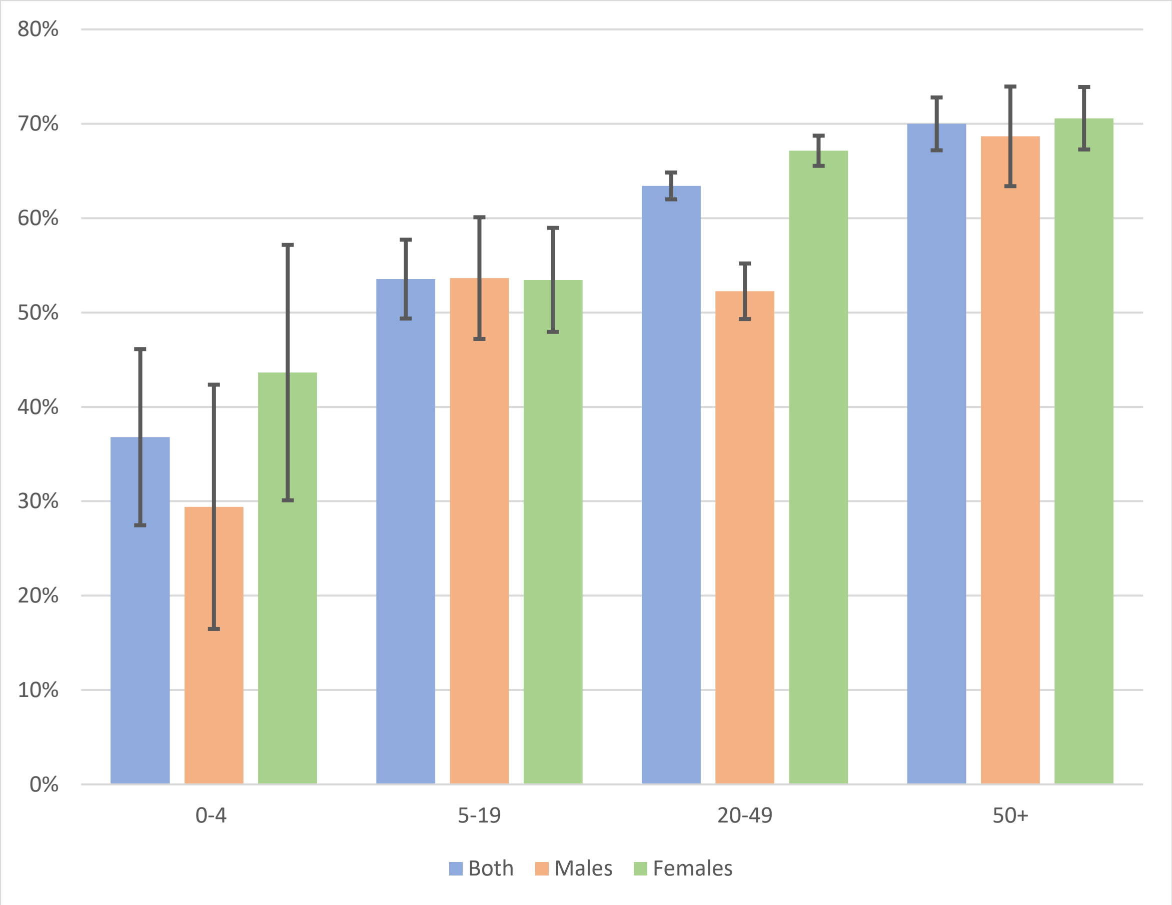


Source: Own calculations from SABSSM 2017 dataset

Indeed, this impact on viral load suppression is patently evident in the SABSSM data. Figure 4, below, provides estimates from SABSSM on the proportion of people living with HIV and exposed to ART who were virally suppressed – defined as a viral load threshold of <1 000 copies HIV RNA/mL – at the time of the survey, disaggregated by age cohort. As is evident in Figure 4, rates of viral load suppression are akin to the aforementioned rates of ART exposure, with rates of viral load suppression highest among the oldest cohort (aged 50 years and over), and lowest among the younger child (aged 0-4 years) and children of school-going age and adolescent (aged 5 – 19 years) cohorts. Specifically, rates of viral load suppression are lowest within the younger child, and children of school-going age and adolescent cohorts, with only

36.8% and 53.5% of individuals living with HIV, respectively, being virally suppressed at the time of the survey. Again, the only statistically significant gender-based difference in rates of viral load suppression occurs in the working age adult (20-49 years) cohort where, again, rates of viral load suppression are higher among females than males.

Figure 4: Viral load suppression among ART-exposed people living with HIV in South Africa, by gender and age cohort



Source: Own calculations from SABSSM 2017 dataset

These data suggest particularly low rates of viral load suppression among children and adolescents of school-going age currently living with HIV in South Africa. Given the evidence in the literature of the negative impact of HIV on educational attainment in a broader sense, and given the hypothesised negative effect of reduced rates of viral load suppression specifically on educational attainment, these data potentially highlight a concerning trend in the potential future human capital development among youth living with HIV in South Africa. Interrogation of this relationship, and of the potential impact of extant reduced rates of ART exposure on educational attainment among the HIV-infected youth in South Africa is, thus, a critically important contribution to the literature.

6.2 Summary of the SABSSM data

Table 1, below, provides summary statistics of the SABSSM data employed in the analyses reported in this paper, covering the chief HIV indicators relevant to this study – HIV status, ART exposure and viral load suppression; the educational attainment indicators relevant to this study – school attendance, absenteeism, age-for-grade status and rates of matriculation; as well as various explanatory covariates, including variables relating to the individual – age, gender, and race; and variables relating to the household – location, cohabitation with parents, single and double orphanhood, household size, and age of household head. Given the focus of the study, specifically the impacts of HIV-related indicators on educational attainment and the potentially gendered nature thereof, these summary statistics are disaggregated by both gender and by HIV status.

Table 1: Summary statistics of the SABSSM dataset for children and adolescents of school-going age in South Africa

	Total			HIV+		
	Both	**Males**	**Females**	**Both**	**Males**	**Females**
	N=20,675	*N=10,145*	*N=10,530*	*N=551*	*N=233*	*N=318*
	Mean (Std Error)	Mean (Std Error)	Mean (Std Error)	Mean (Std Error)	Mean (Std Error)	Mean (Std Error)
HIV positive	0.0266	0.0229	0.0301$^{\#}$			
	(0.0011)	(0.0014)	(0.0016)			
ART exposure				0.3666	0.3304	0.3930
				(0.0205)	(0.0308)	(0.0274)
Viral load suppression				0.5353	0.5364	0.5345
				(0.0212)	(0.0327)	(0.0280)
Education characteristics						
At school	0.7953	0.7930	0.7998	0.7622	0.8369	0.7075$^{\#+}$
	(0.0028)	(0.0040)	(0.0038)	(0.0181)	(0.0252)	(0.0255)
Old for grade[a]	0.1974	0.2266	0.1702$^{\#}$	0.2617^{+}	0.2352	0.2834
	(0.0037)	(0.0054)	(0.0052)	(0.0239)	(0.0351)	(0.0327)
Enforced absenteeism[a]	0.1709	0.1742	0.1678	0.2182^{+}	0.2094	0.2266
	(0.0028)	(0.0041)	(0.0040)	(0.0197)	(0.0289)	(0.0269)
Matriculated[b]	0.1345	0.1155	0.1526$^{\#}$	0.1333	0.1458	0.1282
	(0.0077)	(0.0107)	(0.0112)	(0.0350)	(0.0808)	(0.0384)
Individual characteristics						
Age	11.6608	11.5781	11.7401	13.0871^{+}	12.3133^{+}	13.6540$^{\#+}$
	(0.0298)	(0.0426)	(0.0418)	(0.1866)	(0.2661)	(0.2536)
Female	0.5093			0.5771^{+}		
	(0.0034)			(0.0210)		
African	0.7838	0.7848	0.7829	0.9347^{+}	0.9305^{+}	0.9379^{+}
	(0.0029)	(0.0042)	(0.0041)	(0.0109)	(0.0173)	(0.0141)
Coloured	0.1143	0.1142	0.1143	0.0415^{+}	0.0277^{+}	0.0517^{+}
	(0.0023)	(0.0032)	(0.0032)	(0.0088)	(0.0112)	(0.0130)
Indian	0.0470	0.0468	0.0471	0.0039^{+}	0.0046^{+}	0.0034^{+}
	(0.0015)	(0.0021)	(0.0021)	(0.0027)	(0.0046)	(0.0034)
White	0.0547	0.0540	0.0554	0.0197^{+}	0.0370^{+}	0.0068^{+}
	(0.0016)	(0.0023)	(0.0023)	(0.0045)	(0.0100)	(0.0046)
Household characteristics						
Urban	0.4807	0.4801	0.4811	0.3593^{+}	0.3218^{+}	0.3867^{+}
	(0.0034)	(0.0049)	(0.0048)	(0.0204)	(0.0306)	(0.0273)
Maternal/Paternal Orphan	0.1501	0.1436	0.1567	0.3139^{+}	0.3004^{+}	0.3238^{+}
	(0.0024)	(0.0034)	(0.0035)	(0.0197)	(0.0300)	(0.0262)
Double Orphan	0.0245	0.0243	0.0248	0.0834^{+}	0.0643^{+}	0.0974^{+}
	(0.0010)	(0.0015)	(0.0015)	(0.0117)	(0.0161)	(0.0166)
Household size	6.8116	6.7599	6.8615	7.3206	7.7488	7.0065
	(0.0310)	(0.0457)	(0.0420)	(0.2156)	(0.4116)	(0.2194)
Age of household head	51.4814	51.4883	51.4767	52.4307	54.1928	51.1381
	(0.1180)	(0.1698)	(0.1642)	(0.7038)	(1.0760)	(0.9248)
Household wealth quintile						
Poorest	0.2513	0.2516	0.2495	0.2704	0.2188	0.3081
	(0.0030)	(0.0043)	(0.0042)	(0.0189)	(0.0271)	(0.0259)
Second	0.1924	0.1940	0.1914	0.2141	0.2145	0.2138
	(0.0027)	(0.0039)	(0.0038)	(0.0174)	(0.0269)	(0.0230)
Middle	0.1843	0.1823	0.1866	0.1960	0.2060	0.1886
	(0.0026)	(0.0038)	(0.0037)	(0.0169)	(0.0265)	(0.0219)
Fourth	0.2023	0.1992	0.2057	0.2014	0.2060	0.1981
	(0.0027)	(0.0039)	(0.0039)	(0.0171)	(0.0265)	(0.0223)
Richest	0.1694	0.1726	0.1665	0.1179^{+}	0.1545	0.0911^{+}
	(0.0026)	(0.0037)	(0.0036)	(0.0137)	(0.0237)	(0.0161)

Source: Own calculations from SABSSM 2017 dataset

[a] *Restricted to children and adolescents that are in school*

[b] *Restricted to adolescents of possible matriculation age (17-19 years)*

$^{\#}$ *Statistically significant difference between means for males and females*

$^{+}$ *Statistically significant difference between means for sample population and HIV-positive population*

Approximately 2.7% of the sample of children and adolescents of school-going age (5-19 years) are HIV-positive, with higher rates among females (3.0%) than among males (2.3%), as highlighted above. The rate of ART exposure among the HIV-positive cohort in the sample is 36.7%, with 33.0% of HIV-positive boys and 39.3% of HIV-positive girls in the sample being exposed to ART. The rates of viral load suppression, meanwhile, are more or less equal across gender, with 53.6% of ART-exposed HIV-positive boys being virally suppressed, relative to 53.5% of ART-exposed HIV-positive girls. The four educational attainment indicators that are used in this study are, for all children and adolescents of school-going age (aged 5-19 years), school attendance, absenteeism and age-for-grade status, and, for adolescents of possible matriculation age (17-19 years), matriculation status. All four of these variables are represented by a generated dummy variable. Approximately 79.5% of children and adolescents of school-going age in the sample are attending school, with boys (79.3%) and girls (80.0%) having almost identical attendance rates. The school attendance rate of HIV-positive boys, at 83.74%, is, interestingly, higher than among girls, at 70.8%. Notably, this latter difference in means is statistically significant, both by HIV-positive status (mean attendance rates are lower among HIV-positive girls than among the overall population of girls), and between genders (mean attendance rates are lower among HIV-positive girls than among HIV-positive boys) which already implies a gendered effect of HIV status on school attendance.

Potential evidence of similar gendered impacts is also observable across the other educational attainment dummies under observation: while boys (22.7%) are more likely than girls (17.0%) to be old for their grade, with a statistically significant difference existing between these means, among HIV-positive cohorts specifically, there is no statistically significant difference between these means, implying that HIV-positive status may contribute to an increased likelihood of being old for ones grade among girls specifically. Further evidence of a potentially statistically significant impact of HIV-status on age-for-grade status lies in the statistically significant difference between the mean likelihood of being old for grade among the entire sample (19.7%), relative to among HIV-positive children and adolescents of school-going age specifically (26.2%).

Finally, rates of absenteeism from school are 17.1%, with a fairly even split between males (17.4%) and females (16.8%). Rates of absenteeism are higher among the HIV-positive cohort, at 21.8%, and statistically significant, with these higher rates attributable to higher rates of absenteeism among both HIV-positive boys and HIV-positive girls relative to the overall sample, albeit these differences are not statistically significant by gender.

In terms of the matriculation status of adolescents of possible matriculation age, approximately 13.5% of all such adolescents reported having completed matric or higher, with higher rates among female adolescents of possible matriculation age (15.3%) than among their male counterparts (11.6%), with these differences in means being statistically significant. Among HIV-positive adolescents of possible matriculation age, meanwhile, the overall matriculation rate is lower, albeit not significantly so, with only 13.3% of HIV-positive adolescents of possible matriculation age reporting having completed matric or higher. Notably, among HIV-positive adolescents, the rate of matriculation is higher among HIV-positive males (14.6%) than among HIV-positive females (12.8%), despite the inverse situation across the entire sample of adolescents. Moreover, akin to age-for-grade status, the statistically significant difference between the means along gender lines that exists among the entire sample falls away when the sample is restricted to HIV-positive adolescents of possible matriculation age only, implying that HIV-positive status may contribute to a statistically significant decreased likelihood of having matriculated among girls specifically.

These data imply that there is an overall impact of HIV-status on educational attainment, and that this impact may well be gendered in nature. This calls for further interrogation of these impacts using appropriate statistical and econometric methods.

6.3 Logit regressions

As an initial interrogation of these potential impacts, the association between the relevant educational attainment and HIV-related indicators were estimated using a series of logit regressions and, subsequently, through the measurement of the marginal effects of the latter on the former. For each of the 12 potential relationships (one each between the four indicators of educational attainment and three indicators along the HIV treatment cascade), a set of four logit models were estimated: the first model interrogates the relationship between the relevant HIV indicator and the relevant educational attainment indicator in the absence of any individual or household covariates; the second model introduces these covariates (as presented in Table 1 above); the third model removes these covariates, but introduces an HIV-gender interaction term; and the final model includes both the HIV-gender interaction term and the individual and household covariates. Subsequent to these logit estimations, the marginal effects of the HIV indicators on the educational attainment indicators were estimated using post-estimation techniques. A summary of the logit marginal effects results is presented in Table 2.

Table 2: HIV/AIDS and educational attainment among children and adolescents of school-going age in South Africa

	Interaction term included	Covariates included	HIV positive		ART exposure		Viral load suppression	
			HIV-positive	HIV-positive *Female	ART exposure	ART exposure *Female	VL suppression	VL suppression *Female
At school	No	No	-.1982 [-.0340]		-.0419 [-.0076]		.1661 [.0301]	
	No	Yes	-.2715** [-.0414**]		.0429 [.0058]		.1487 [.0204]	
	Yes	No	.2652 [.0397]	-.7519** [-.1471**]	.6112 [.1047]	-.9634** [-.1943**]	.7710** [.1389**]	-.9449** [-.1830**]
	Yes	Yes	.3627 [.0453]	-1.0098** [-.1879**]	.3970 [.0524]	-.4924 [-.0734]	.2607 [.0358]	-.1590 [-.0222]
Old for grade	No	No	.3764** [.0660**]		.0011 [.0002]		-.2295 [-.0445]	
	No	Yes	.2613** [.0419**]		-.1294 [-.0247]		-.1974 [-.0382]	
	Yes	No	.2346 [.0396]	.2511 [.0427]	-.0866 [-.0166]	.1493 [.0294]	-.3876 [-.0754]	.2747 [.0542]
	Yes	Yes	-.0173 [-.0025]	.5044 [.0866]	-.0822 [-.0157]	-.0796 [-.0151]	-.1343 [-.0259]	-.1085 [-.0207]
Absenteeism	No	No	.2881** [.0447**]		.7472** [.1321**]		.1378 [.0234]	
	No	Yes	.3179** [.0487**]		.6967** [.1184**]		.2844 [.0463]	
	Yes	No	.2310 [.0352]	.1072 [.0158]	.7189** [.1269**]	.0512 [.0086]	-.1066 [-.0181]	.4564 [.0819]
	Yes	Yes	.2470 [.0370]	.1334 [.0193]	.6216 [.1051]	.1404 [.0232]	-.2136 [-.0349]	.9396 [.1721]
Matriculated	No	No	-.0105 [-.0012]		-.1288 [-.0147]		.2516 [.0539]	
	No	Yes	-.4659 [-.0351]		.5693 [.0375]		.2017 [.0435]	
	Yes	No	.0910 [-.1492]	.0109 [-.0165]	-.3801 [-.0430]	.3364 [.0405]	.2231 [.0478]	.0358 [.0077]
	Yes	Yes	.2852 [.0282]	-1.008 [-.0621**]	.0593 [.0037]	.7291 [.0519]	-.0353 [-.0076]	.3011 [.0656]

Source: Own calculations from SABSSM 2017 dataset. Marginal effects of logit regression models. Children and adolescents represent respondents between the age of 5 and 19 years old.
*** p<0.05*
Marginal effects in square brackets
Absenteeism and age-for-grade estimations restricted to children of school-going age and adolescents in school
Matriculation estimations restricted to adolescents of possible matriculation age (17-19 years)

As is evident in Table 2, there is a clear statistically significant relationship between HIV-positive status and three chief educational outcome indicators, namely: school attendance, absenteeism, and age-for-grade status, when controlling for the impact of the relevant individual and household covariates. Specifically, the HIV-positive coefficient is negative and statistically significant at the 5% level in the school attendance model, which includes the individual and household covariates, which indicates that being HIV-positive is statistically significantly associated with a decreased likelihood of being at school among children and adolescents of school-going age in South Africa. Similarly, the HIV-positive coefficient is positive and statistically significant at the 5% level in both the age-for-grade status and absenteeism models when the individual and household covariates are included. This suggests that being HIV-positive is statistically significantly associated with an increased likelihood of being absent from school, and being old for one's grade, among children and adolescents of school-going age in South Africa that are currently in school. Finally, the ART-exposure coefficient is positive and statistically significant at the 5% level in the school absenteeism model with included covariates, which suggests that ART exposure is positively associated with rates of absenteeism among children and adolescents of school-going age in South Africa, suggesting that ART exposure is statistically significantly associated with an increased likelihood of being absent from school.

Under the fourth set of models, which introduce an HIV indicator-gender interaction term, to interrogate the potential contribution of gender to the modelled impacts, the coefficient of the HIV-related indicators are no longer statistically significant in any of the models. Instead, the coefficient of the interaction term is statistically significant in two of the modelled relationships – between HIV-positive status and school attendance, and between HIV-positive status and matriculation status. This implies that gender, specifically, may contribute to the impacts of HIV-status on indicators of educational attainment, and warrants further interrogation.

6.4 Propensity score matching analysis

The effect of the various HIV indicators on the aforementioned indicators of educational attainment are further interrogated using propensity score matching-based approaches.

Propensity scores were generated using indicators of HIV status and the individual and household characteristics (as outlined in Table 1).

Table 3: Quality of matching in propensity score matching

	LR chi^2	p>chi^2	MeanBias	B
HIV variable: HIV-positive; outcome: school attendance				
Unmatched	207.44	0.000	24.9	65.8*
Matched	0.00	1.000	0.0	0.0
HIV variable: HIV-positive; outcome: absenteeism				
Unmatched	166.91	0.000	23.5	69.0*
Matched	0.00	1.000	0.0	0.0
HIV variable: HIV-positive; outcome: old for grade				
Unmatched	142.20	0.000	25.1	69.8*
Matched	0.00	1.000	0.0	0.0
HIV variable: HIV-positive; outcome: matriculation				
Unmatched	207.44	0.000	24.9	65.8*
Matched	0.00	1.000	0.0	0.0
HIV variable: ART exposure; outcome: school attendance				
Unmatched	7.40	0.388	8.1	25.1*
Matched	0.13	1.000	0.7	3.7
HIV variable: ART exposure; outcome: absenteeism				
Unmatched	6.24	0.512	9.0	26.5*
Matched	0.02	1.000	0.2	1.5
HIV variable: ART exposure; outcome: old for grade				
Unmatched	7.79	0.352	11.4	32.7*
Matched	0.06	1.000	0.7	3.3
HIV variable: ART exposure; outcome: matriculation				
Unmatched	7.40	0.388	8.1	25.1*
Matched	0.13	1.000	0.7	3.7
HIV variable: viral load suppression; outcome: school attendance				
Unmatched	8.23	0.313	7.6	25.2*
Matched	0.18	1.000	1.0	3.7
HIV variable: viral load suppression; outcome: absenteeism				
Unmatched	10.01	0.188	12.2	31.4*
Matched	0.17	1.000	2.9	4.1
HIV variable: viral load suppression; outcome: old for grade				
Unmatched	9.79	0.201	10.2	32.8*
Matched	0.78	0.998	2.3	9.4
HIV variable: viral load suppression; outcome: matriculation				
Unmatched	8.23	0.313	7.6	25.2*
Matched	0.18	1.000	1.0	3.7

Source: Own calculations from SABSSM 2017 dataset

Three commonly used matching algorithms – nearest neighbour, stratification and kernel matching – were employed to identify the nearness between control and treated cases in each instance, and, subsequently, to interrogate the average treatment effect on the treated. In order to ensure the veracity and robustness of the propensity score matching approaches used and the subsequent results, a series of tests were run pre- and post-matching for each of the propensity score matching models used in the study. For all matching algorithms, tests were used to ensure that all covariates were balanced, and that differences between all covariates were insignificant subsequent to matching. Moreover, for each matching algorithm, the chi^2 and a Rubin's B statistics were used to ensure statistical insignificance of the overall model and overall balance respectively. Finally, for each matching algorithm employed, large reductions in both the likelihood ratio (LR) chi^2 statistic and the means-standardized bias were observed post-matching, indicative of a robust matching algorithm. A summary of these pre- and post-matching tests is provided in Table 3 above. Further to the above, all observations for the propensity score matching models used in the analyses below have common support.

Table 4, below, presents the summary results of the 36 propensity score matching models (three HIV indicators, by four educational attainment indicators, by three matching algorithms) as applied to children and adolescents of school-going age, for impacts on school attendance, absenteeism and age-for-grade status; and to adolescents of possible matriculation age, for impacts on matriculation status.

Table 4: Average treatment effect of HIV indicators on indicators of educational attainment, children and adolescents of school-going age

	Matching algorithm		
	Nearest neighbour	**Stratification**	**Kernel**
	HIV-positive		
At school	-.0509**	-.0484**	-.0280**
Old for grade	.0713**	.0613**	.0805**
Absenteeism	.0517**	.0533**	.0416**
Matriculated	-.0269	-.0223	-.0253
	ART exposure		
At school	.0094	-.0014	-.0071
Old for grade	.0018	-.0121	-.0159
Absenteeism	.1028**	.0958**	.1006**
Matriculated	.0210	.0090	.0169
	Viral load suppression		
At school	.0116	.0008	.0026
Old for grade	-.0302	-.0378	-.0448
Absenteeism	.0240	.0200	.0272
Matriculated	.0195	-.0120	-.0345

Source: Own calculations from SABSSM 2017 dataset.
*** p<0.05*
Absenteeism and age-for-grade estimations restricted to children of school-going age and adolescents in school
Matriculation estimations restricted to adolescents of possible matriculation age (17-19 years)

The results in Table 4 suggest three clear statistically significant treatment effects of HIV-related indicators on indicators of educational attainment among children and adolescents of school-going age – namely, of HIV-positive status on school attendance; of HIV-positive status on age-for-grade status; and of HIV-positive status on absenteeism – with the coefficients for these effects being statistically significant across all three matching algorithms employed. The coefficient for HIV-positive status on school attendance is negative and statistically significant across nearest neighbour (-0.051), stratification (-0.048) and kernel (-0.028) matching approaches, which indicates that, among children and adolescents of school-going age, being HIV-positive is negatively associated with school attendance, and that this association is statistically significant. The coefficient for HIV-positive status on age-for-grade, meanwhile, is positive and statistically significant across nearest neighbour (0.071), stratification (0.061) and kernel (0.081) matching approaches, which indicates that, among children and adolescents of school-going age that are in school, being HIV-positive is positively associated with being old for one’s grade, and that this association is statistically significant. The coefficient for HIV-positive status on absenteeism, meanwhile, is positive and statistically significant across nearest neighbour (0.052), stratification (0.053) and kernel (0.042) matching approaches which indicates that, among children and adolescents of school-going age that are in school, being HIV-positive is positively associated with the likelihood of being absent from school, and that this association is statistically significant. The lack of statistically significant results across the other interrogated relationships between HIV-positive status and educational attainment suggests that the status of being HIV-positive is not statistically significantly associated with the likelihood of having matriculated.

When interrogating the average treatment effect of the subsequent HIV indicators on the 90-90-90 treatment cascade – specifically ART exposure and viral load suppression – on the aforementioned indicators of educational attainment, none of the models predicted a statistically significant treatment effect, with the lone exception of the impact of ART exposure on absenteeism under the stratification matching approach. This lone statistically significant result suggests that ART exposure is positively associated with rates of absenteeism among school-going aged children, and supports the seemingly contradictory positive relationship between ART exposure and absenteeism identified in the initial logit estimations.

Subsequent to a propensity score matching-based analysis of the average treatment effect of these HIV indicators on indicators of educational attainment across the entire population using the aforementioned models and matching algorithms, the approach was repeated twice – once with the population restricted to females, and again with the population restricted to males – using the same approach to the generation propensity scores and the same matching algorithms, in order to interrogate the average treatment effect on the treated for each gender separately. Table 5 presents the results of the female-only models, while Table 6 presents the results of the male-only models. The results of the female-only propensity score models closely mirror those of the overall models as summarized in Table 4, previously, with the average treatment effect of HIV status on school attendance, age-for-grade status, and absenteeism status similarly identified as being statistically significant. Again, these effects are identified in the same direction as in the overall models – a positive effect of HIV-positive status on being old for one's grade and the likelihood of absenteeism, and a negative effect of HIV-positive status on the likelihood of being in school – but the size of the coefficient is larger in all three instances, suggesting a larger effect. Unlike in the overall model, in the female-only models there exists a statistically significant negative average treatment effect of HIV-positive status on matriculation status, but only under two of the three matching algorithms – nearest neighbour and stratification matching. Both of these statistically significant results are negative, implying a negative effect of HIV-positive status on the likelihood of matriculation among female adolescents of possible matriculation age.

Table 5: Average treatment effect of HIV indicators on indicators of educational attainment, female children and adolescents of school-going age

	Matching algorithm		
	Nearest neighbour	**Stratification**	**Kernel**
	HIV-positive		
At school	-.1140**	-.1108**	-.0921**
Old for grade	.1209**	.1218**	.1326**
Absenteeism	.0637**	.0675**	.0610**
Matriculated	-.0447**	-.0473**	.0079
	ART exposure		
At school	-.0032	.0051	-.0098
Old for grade	-.0020	-.0028	-.0229
Absenteeism	.0997**	.0833	.1090**
Matriculated	.0266	.0191	.0262
	Viral load suppression		
At school	-.0248	-.0355	-.0366
Old for grade	-.0280	-	-.0296
Absenteeism	.1019	.0921	.0997
Matriculated	.0423	.0454	.0181

Source: Own calculations from SABSSM 2017 dataset
*** $p<0.05$*

Absenteeism and age-for-grade estimations restricted to children of school-going age and adolescents in school
Matriculation estimations restricted to adolescents of possible matriculation age (17-19 years)

Table 6: Average treatment effect of HIV indicators on indicators of educational attainment, male children and adolescents of school-going age

	Matching algorithm		
	Nearest neighbour	**Stratification**	**Kernel**
	HIV-positive		
At school	.0338	.0338	.0515
Old for grade	.0081	-.0128	.0181
Absenteeism	.0389	.0361	.0242
Matriculated	.0172	.0181	.0182
	ART exposure		
At school	.0296	.0221	.0236
Old for grade	.0055	-.0292	-.0091
Absenteeism	.1033	.0792	.0867
Matriculated	.0069	-.0170	-.0096
	Viral load suppression		
At school	.0602	.0510	.0546
Old for grade	-.0292	-.0499	-.0650
Absenteeism	-.0633	-.0425	-.0497
Matriculated	-.0649	-.1190	-.1724

Source: Own calculations from SABSSM 2017 dataset
*** $p<0.05$*
Absenteeism and age-for-grade estimations restricted to children of school-going age and adolescents in school
Matriculation estimations restricted to adolescents of possible matriculation age (17-19 years)

The results of the male-only propensity score models, meanwhile, suggest that there is no statistically significant effect of any of the HIV-related indicators on any of the indicators of educational attainment. As is illustrated in Table 6, this lack of any statistically significant average treatment effects suggests that, among male children and adolescents of school-going age in South Africa, HIV status, ART exposure, and viral load suppression have no effect on the likelihood of being out of school, being absent from school, being old for one's grade or having completed matric. The results of these gender-specific propensity score models suggests a gendered impact of HIV-related indicators on indicators of educational attainment among children and adolescents of school-going age in South Africa, specifically that HIV-related factors have an impact on educational attainment for girls that does not exist for boys. The gendered nature of this impact warrants further investigation.

6.5 Oaxaca-Blinder decomposition analysis

Subsequent to the interrogation of the average treatment effects of various HIV-related indicators on educational attainment through logit and propensity score-based estimations, and an identification of the gendered nature of these effects, it is important to understand the extent to which gender contributes to gender gaps in educational attainment by decomposing the identified gaps in educational attainment. Decomposition analyses were conducted to interrogate how the various HIV-related factors influence the gender gap in overall educational attainment outcomes among all children and adolescents of school-going age in South Africa. The results of these decompositions are presented in Table 7, below.

Table 7: Detailed Oaxaca-Blinder decomposition analyses results, by gender

	Outcome Variables							
	At school		Absenteeism		Old for grade		Matriculated	
Male	**.8236**		**.1742**		**.2294**		**.1118**	
Female	**.8298**		**.1677**		**.1737**		**.1515**	
Difference (Male - Female)	**-.0062**		**.0065**		**.0556****		**-.0396****	
Of which	***Explained Differential***	***Unexplained Differential***	***Explained Differential***	***Unexplained Differential***	***Explained Differential***	***Unexplained Differential***	***Explained Differential***	***Unexplained Differential***
HIV cascade indicators								
HIV-positive	.0003	.0030**	.0000	.0014	-.0002	-.0029	.0011	.0055
ART exposure	.0001	.0000	-.0001	.0001	-.0000	-.0000	-.0013	.0009
Viral load suppression	-.0001	.0005	.0000	-.0019	.0000	.0002	.0024	-.0034
Individual characteristics								
Age	-.0000	.0258	-.0002	-.0209	.0005	.1392**	-.0040	-.1793
African	.0000	-.0158	-.0000	-.0104	-.0000	.0414	-.0000	.0194
Coloured	.0000	.0004	-.0000	-.0011	-.0000	.0071	-.0001	.0057
Indian	-.0000	.0021	.0003	-.0031	-.0003	.0014	.0006	.0030
Household characteristics								
Urban	.0000	-.0104	-.0003	.0014	.0000	-.0041	.0003	-.0158
Both parents in household	.0006	-.0039	-.0000	.0023	.0012	.0074	-.0003	.0100
Maternal/Paternal Orphan	-.0007	.0001	-.0002	-.0006	-.0006	-.0017	.0000	-.0026
Double Orphan	.0000	-.0001	.0000	-.0005	.0000	.0032	.0000	.0000
Household size	.0001	-.0169	.0001	.0181	-.0001	-.0202	-.0000	-.0306
Age of household head	.0000	-.0015	.0000	.0044	.0000	.0303	-.0000	-.0302
Household wealth quintile								
Poorest	-.0001	-.0079	.0000	-.0009	-.0002	.0023	-.0002	-.0005
Second	-.0000	.0003	-.0000	.0006	-.0000	.0056	-.0004	-.0046
Middle	.0000	-.0030	.0000	-.0047	-.0002	.0036	.0003	.0012
Fourth	.0000	.0002	.0000	-.0024	.0000	-.0023	.0000	-.0055
Total	**.0003**	**-.0065**	**-.0007**	**.0072**	**-.0000**	**.0557****	**-.0017**	**-.0379****

Source: Own calculations from SABSSM 2017 dataset
*** $p<0.05$*
Absenteeism and age-for-grade estimations restricted to children of school-going age and adolescents in school
Matriculation estimations restricted to adolescents of possible matriculation age (17-19 years)

As is evident in Table 7, and was discussed in the summary analyses earlier, among the entire sample of children and adolescents of school-going age, there are statistically significant gender gaps for age-for-grade status and rates of matriculation – with girls being less likely to be old for their grade and more likely to have matriculated. Decomposition of these gender gaps suggests that none of the HIV-related factors contribute significantly to the observed gender gaps. Indeed, of the included observed characteristics in both decompositions, only one is a statistically significant contributor to the observed gender gap – namely age, which contributes in a statistically significant manner to the age-for-grade gender gap. What is notable in the overall decompositions summarized in Table 7, however, is the fact that, despite there not being a statistically significant gender gap in school attendance among children and adolescents of school-going age, under decomposition, the unexplained impact of HIV-positive status on gender gaps in school attendance is statistically significant, and positive. This suggests that there is an unexplained pro-boy gender gap in the impact of HIV-positive status on school attendance, meaning that the status of being HIV-positive has a statistically significant impact on the likelihood of girls not being in school.

In order to interrogate the nature of gender gaps in educational attainment outcomes among children and adolescents of school-going age on the HIV treatment cascade, specifically, a further set of decomposition analyses were conducted in order to decompose gaps in educational attainment by gender along the HIV treatment cascade, by conducting separate gender-based decompositions within sub-samples along this cascade. Such decomposition allows for an examination of the explained and unexplained differential in educational attainment among HIV status cohorts *across* genders. These decompositions separate the identified gender gaps in educational attainment into an explained differential – attributable to the characteristics covered by the covariates included in the model; and an unexplained differential – attributable to factors that are not observed in the data, and which are typically assumed to be indicative of a gender gap. Table 8, below, summarizes the results of these decompositions, showing the difference in coefficients between males and females for each of the relevant educational attainment outcomes, and, subsequently, the contribution to said differences in coefficients that are the effects of differences in endowments (explained differential), and the effects of differences in coefficients (unexplained differential).

Table 8: Oaxaca-Blinder decomposition analysis results, within HIV-status, by gender

		HIV-positive	ART exposure	Viral load suppression
At school	Difference (Male - Female)	.1316**	.1462**	.1457**
	Explained	*.0123*	*-.0082*	*.0325*
	Unexplained	*.1192***	*.1545***	*.1131***
Old for grade	Difference (Male - Female)	-.0598	-.0498	-.0737
	Explained	*-.0520*	*-.0493*	*-.0409*
	Unexplained	*-.0078*	*-.0005*	*-.0327*
Absenteeism	Difference (Male - Female)	-.0246	-.0485	-.1012
	Explained	*-.0129*	*-.0501*	*-.0213*
	Unexplained	*-.0116*	*.0016*	*-.0799*
Matriculated	Difference (Male - Female)	.0248	-.0014	-
	Explained	*-.0686*	*.0146*	-
	Unexplained	*.0934*	*-.0160*	-

Source: Own calculations from SABSSM 2017 dataset
** $p<0.05$

Absenteeism and age-for-grade estimations restricted to children of school-going age and adolescents in school
Matriculation estimations restricted to adolescents of possible matriculation age (17-19 years)

As is evident in Table 8, there is a statistically significant difference in rates of school attendance between males and female children and adolescents of school-going age across all three of the HIV treatment cascade indicators, which is indicative of a notable gender gap in school attendance along the HIV treatment cascade. For all three of the indicators on this cascade, the unexplained differential represents the sole statistically significant contributor to these gender gaps, which suggests that the identified gender gaps are attributable to gender itself, rather than to other individual and household characteristics related to gender. Given the statistically significant positive differential between coefficients, which implies higher rates of school attendance among HIV-positive, ART-exposed and virally suppressed boys than their girl counterparts, the statistical significance of the unexplained differential in all three instances suggests that HIV-positive status, ART exposure, and viral load suppression has a pro-female impact on school non-attendance that is unexplained by factors beyond gender.

An additional set of decomposition analyses were conducted to interrogate the impact of various indicators to gaps in educational attainment among female children and adolescents of school-going age specifically, in order to identify the factors that contribute to gaps in educational attainment among females along the HIV treatment cascade. An initial set of decompositions were run to identify statistically significant unexplained differences within the various relationships. Subsequently, a detailed decomposition was run for each statistically significant unexplained difference in order to examine the factors that contribute to the identified unexplained statistically significant gaps. Table 9 presents the summary results of the initial full set of decompositions, while presents the results of the subsequent detailed decompositions.

Table 9: Summary of Oaxaca-Blinder decomposition analysis results, within females, by HIV status

		At school	Old for grade	Absenteeism	Matriculation
HIV-positive	Difference (HIV negative - HIV positive)	0.0841**	-0.0695	-0.0417	0.0812**
	Endowments	*-0.0084*	*0.0395*	*0.0020*	*0.0420*
	Coefficients	*0.1023***	*-0.1294***	*-0.0480*	*0.0903*
	Interaction	*-0.0097*	*0.0204*	*0.0042*	*-0.0512*
ART exposure	Difference (ART naïve - ART exposed)	0.0119	0.0223	-0.1251**	-0.0088
	Endowments	*0.0005*	*0.0223*	*-0.0314*	*-0.0017*
	Coefficients	*0.0104*	*-0.0061*	*-0.1160*	*-0.0018*
	Interaction	*0.0009*	*0.0061*	*0.0223*	*-0.0052*
Viral load suppression	Difference (VL unsuppressed - ART suppressed)	-0.0181	0.0986	-0.0961	-0.0602
	Endowments	*-0.0061*	*0.0130*	*0.0073*	*0.0063*
	Coefficients	*-0.0095*	*0.0630*	*-0.1129*	*-0.0609*
	Interaction	*-0.0024*	*0.0225*	*0.0094*	*-0.0056*

Source: Own calculations from SABSSM 2017 dataset
*** p<0.05*
Absenteeism and age-for-grade estimations restricted to children of school-going age and adolescents in school
Matriculation estimations restricted to adolescents of possible matriculation age (17-19 years)

Table 9 identifies, for females, similar statistically significant impacts of HIV indicators on educational attainment as identified in the propensity score matching- and logit-based approaches – namely a statistically significant impact of HIV-positive status on school attendance and age-for-grade status – evidenced by the statistically significant differences in both in Table 9. Further, it is notable that the unexplained difference – which would be the difference assumed to be attributable to the relevant HIV indicator, rather than to some other factors – is statistically significant for the effect of HIV-positive status on all four indicators of educational attainment, suggesting that HIV-positive status has a significant unexplained impact on the likelihood of female children and adolescents of school-going age being out of school, old for their grade, absent from school, and having matriculated. A final statistically significant impact within the female decomposition is of ART exposure on absenteeism, which mirrors the finding under the propensity score matching model as summarized in Table 5.

Table 10: Detailed unexplained coefficients for statistically significant Oaxaca-Blinder decomposition analysis results, within females, by HIV status

Outcome	At school	Absenteeism	Old for grade	Matriculated	Absenteeism
By	**HIV-positive**	**HIV-positive**	**HIV-positive**	**HIV-positive**	**ART exposure**
Unexplained Difference	**.1009****	**-.0619****	**-.0968****	**.0724****	**-.1270****
Individual characteristics					
Age	.2262**	-.1644	-.1558	.7632	-.7918**
African	.4638**	.1345	.1131	-.3561**	.2124
Coloured	.0263**	-.0150**	.0228	-.0202	.0243
Indian	.0039	-.0076**	.0050	-.0018	.0026
Household characteristics					
Urban	-.0027	.0047	-.0404	-.0277	-.0596
Both parents in household	.0126	.0042	-.0117	-.0314	.0495
Maternal/Paternal Orphan	-.0506**	.0390	.0164	.0058	.0405
Double Orphan	-.0001	-.0031	-.0189	-.0228	.0207
Household size	.0694	.0254	-.1089	.0062	-.0679
Age of household head	-.2114	-.2530	.0894	.0969	.1263
Household wealth quintile					
Poorest	-.0076	-.0133	-.0179	-.0268	-.0031
Second	.0185	-.0047	.0156	-.0375	.0317
Middle	-.0042	-.0206	.0028	-.0191	.0078
Fourth	-.0156	-.0107	-.0053	.0088	-.0226

Source: Own calculations from SABSSM 2017 dataset
*** $p<0.05$*
Absenteeism and age-for-grade estimations restricted to children of school-going age and adolescents in school
Matriculation estimations restricted to adolescents of possible matriculation age (17-19 years)

Table 10 shows that, across the various unexplained statistically significant differences in educational attainment outcomes among female children and adolescents of school-going age specifically along the HIV treatment cascade, age and race are the two categories of variable that make a statistically significant contribution to these observed differences, suggesting that there may be unexplained race- and age-based gaps in the impact of HIV-related factors on educational attainment among female children and adolescents of school-going age in South Africa.

Akin to the above female-specific decomposition, a final set of decomposition analyses were conducted to interrogate the impact of various indicators to gaps in educational attainment among male children and adolescents of school-going age specifically, in order to identify the factors that contribute to gaps in educational attainment among males along the HIV treatment cascade. These decompositions found no statistically significant differences within the various relationships, suggesting that HIV-related indicators have no statistically significant impact on indicators of educational attainment for male children and adolescents of school-going age in

South Africa. Consequently, the results of these decompositions are not presented here, but may be provided upon request.

7 Discussion

Recent contributions to the literature on the relationship between HIV and education in children has suggested that HIV contributes to detrimental education outcomes, realized as increased rates of school dropout, increased rates of absenteeism, and poor educational performance and throughput rates (Zinyemba, Pavlova and Groot, 2020). However, there has been scant research to quantify the extent to which HIV might impact on each of these outcomes, and even less research on the potential gendered nature of these impacts, the associated gender gaps in education outcomes, and potential differences in HIV-related gaps in educational outcomes within each gender. Moreover, extant literature on relationship between HIV and education in children has focussed explicitly on HIV-status and has ignored the impact of HIV-related factors further down the HIV treatment cascade – ART exposure and viral load suppression – which are more typically associated with the hypothesized mechanism through which HIV impacts on education outcomes, namely increased rates of illness and morbidity, than mere HIV-positive status. This study has sought to address this gap in the literature by estimating the effects of three chief indicators along the HIV treatment cascade – HIV status, ART exposure and viral load suppression, on four indicators of educational attainment – school attendance, absenteeism, age-for-grade status and matriculation status, among children and adolescents of school-going age (aged 5-19 years) in South Africa. The study has employed data with biomedical information on HIV for children and adolescents in South Africa from the 5th South African National HIV Prevalence, Incidence, Behaviour and Communication Survey (SABSSM V), conducted in 2017/8 by the Human Sciences Research Council, to examine the impact of the aforementioned HIV indicators on educational attainment among HIV children in South Africa. Further to this, the study has also interrogated whether there are gender differences in how these HIV indicators impact on educational attainment, and assessed the different contributions to HIV-related educational attainment gaps within each gender.

The initial analyses find that HIV-positive status specifically has a statistically significant negative impact on the likelihood of being in school among all children and adolescents of school-going age in South Africa. This finding is in line with prior research on the impact of

HIV-status on school attendance which has found that being HIV-positive has a negative impact on school attendance, notably by Parchure *et al*. (2016) which found that, among children aged 6-16 years in Maharashtra, India, HIV-infected children were seven times more likely to be out of school; and Kembo (2010), which found HIV-affected children aged 10-18 years in Zimbabwe to be out of school at a significantly higher rate than HIV-unaffected children. The initial analyses, similarly, find that HIV-positive status has a statistically significant positive impact on the likelihood of absenteeism among all children and adolescents of school-going age in South Africa. Again, this result is in concurrence with studies on the relationship between HIV-status and absenteeism, including Anabwani *et al*. (2016) which found a significant impact of HIV-positive status on absenteeism among children aged 6–18 years in Botswana; and Cohen *et al*. (1997) which found significantly increased rates of absenteeism among HIV-positive children in Massachusetts, USA. Finally, these initial analyses find that HIV-positive status has a statistically significant positive impact on the likelihood of being old for one's grade among all children and adolescents of school-going age in South Africa, in line with studies by Bandason *et al*. (2013), Fotso *et al*. (2018), and Henning *et al*. (2018) which find a similar statistically significant impact of HIV-status on age-for-grade status among children in Zimbabwe, South Africa and Rwanda, respectively.

Analyses of the second HIV treatment cascade indicator – ART exposure – on indicators of educational attainment found one statistically significant impact, namely a positive impact of ART exposure on the likelihood of absenteeism among children and adolescents of school-going age in South Africa. This is a seemingly contradictory finding, as ART exposure is typically associated with HIV viral load suppression and subsequent reductions in HIV-related morbidity, which is the hypothesized mechanism through which HIV-positive status would impact on the likelihood of absenteeism. However, this finding does support findings from other studies in the literature which have found similar positive relationships between ART exposure and absenteeism, albeit for work absenteeism among working-age adults. Specifically, studies by Habyarimana *et al*. (2010), on the impact of antiretroviral treatment on worker absenteeism in Botswana, and Jockers *et al*. (2021), on the impact of ART on worker absenteeism in South Africa, both found robust evidence of an inverted-V-shaped pattern in worker absenteeism around the time of ART inception. This pattern suggests that absenteeism peaks significantly in the 6-12 month period prior to ART inception, and then declines significantly, albeit at a slightly slower rate than the aforementioned increase, before returning to normal levels after a period of around 18 months (Jockers *et al*., 2021; Habyarimana *et al*., 2010). This suggests that ART initiation may be driven by HIV-related morbidity, with HIV-positive individuals that are suffering from morbidity specifically seeking antiretroviral treatment. Moreover, the right-tailed nature of the identified inverted-V-shaped relationship between ART exposure and absenteeism would mean that, averaged over the course of ART initiation, absenteeism rates would be higher during the ART-exposed period. These two factors may explain the apparently contradictory positive relationship between ART exposure and school absenteeism which has been identified here. A further contributing factor may be the impact of ART-related side-effects – with several studies finding that a significant proportion of patients on antiretroviral treatment suffer adverse side-effects, most notably anaemia (Zhou *et al*., 2012), gastrointestinal disorders (Hill and Balkin, 2009), and metabolic and renal morbidities (Fortuny *et al*., 2015) – which may contribute to increased rates of absenteeism.

Finally, the analysis of the final downstream HIV treatment cascade indicator – viral load suppression – on educational attainment found no negative impact of viral non-suppression on any of the indicators of educational attainment among children and adolescents of school-going age in South Africa. This is akin to a study by Souza *et al*. (2010) in Brazil, which, despite finding higher rates of school failure and dropout among HIV-positive children aged 10-19 years, found no relationship between HIV viral load and said education outcomes. This suggests that the previously identified negative impact of HIV-positive status on educational attainment among South African children and adolescents of school-going age may not manifest through the hypothesized mechanism of increased rates of morbidity and illness-related underperformance, absenteeism and dropout among HIV-positive children, as HIV-related morbidity is typically associated with viral non-suppression and it would thus be expected that this downstream indicator would also be statistically significantly associated with the detrimental education outcomes identified in the HIV-status models. Instead, this finding suggests that the status of being HIV-positive may impact on educational attainment through other mechanisms. One potential mechanism is the reduced demand for investment in education among HIV-positive children and adolescents of school-going age, as per the Grossman model of health demand which posits that the poorer a child's health status, the lower the expected return on an investment in their education, and the lower the demand for such an investment. A study by Kitara *et al*. (2013) finds that HIV-affected girls aged 12-15 in

Uganda had a less positive attitude to school, relative to HIV-unaffected girls, which may be the result of the above mechanism. Further mechanisms may include internal and external stigma associated with HIV-status, as identified by Henning *et al*. (2018) in an analysis of age-for-grade status among HIV-positive and HIV-affected children in Rwanda; affective and social impacts of being HIV-positive, as noted by Ellis (2004) in a study of the grade performance of perinatally-infected HIV-positive children in North Carolina; or the negative effects on the cognitive functioning of children and adolescents associated with perinatal HIV infection, as identified by Phillips *et al*. (2016).

The subsequent gender-disaggregated propensity score matching-based analyses interrogated the gendered nature of the impacts of HIV-related factors on educational attainment among children and adolescents of school-going age in South Africa. This initial gender analysis identified several gendered differences in the impact of HIV-related factors on educational attainment – with HIV-positive status having a statistically significant impact on all four indicators of educational attainment under observation for female children and adolescents of school-going age, while having no statistically significant impact for their male counterparts. An initial set of decompositions further interrogated the gender-HIV-education nexus by examining the contribution of HIV-related indicators to observed gender gaps in education overall, and a subsequent set of decompositions examined the nature of gender gaps in educational attainment within each cohort of the HIV treatment cascade. The initial decompositions found only one statistically significant contribution of an HIV-related indicator to gender gaps in educational attainment among the entire sample of children and adolescents of school-going age – namely an unexplained pro-boy gender gap in the impact of HIV-positive status on school attendance. This lack of statistically significant impacts when observing the entire sample is to be expected – there are minimal gender gaps in educational attainment overall, attributable to gender parity in primary and secondary school enrolment in South Africa (World Bank, 2023), and attributable to the small size of the HIV-positive sub-sample – but the aforementioned pro-boy gender gap in the impact of HIV-positive status on school attendance does add further credence to the disproportionate negative impact of being HIV-positive on school attendance on female children and adolescents of school-going age, as identified in the propensity score matching-based analyses. This is further evidenced in the subsequent decompositions, conducted within sub-samples along the HIV treatment cascade. These decompositions found unexplained statistically significant gender gaps in school attendance within all three sub-samples along the HIV treatment cascade, suggesting that observed statistically significant gender gaps in school attendance along the cascade are directly attributable to gender, rather than to the effects of differences in endowments. Together, these findings suggest a female-specific negative impact of HIV-status on educational attainment, similar to Delva *et al*. (2009), who found that, among HIV-affected children aged 10-18 in Guinea, HIV-affected boys were more likely to attend school than their HIV-affected female counterparts; Bhargava (2005), who found that HIV-affected girls in Ethiopia were less likely to participate in school than HIV-affected boys; and Harrison *et al*. (2017), who found that HIV-affected girls aged 6-17 in China reported lower grades and less interest in school than their male counterparts; but distinct from studies by Curley *et al*. (2010), Hensels *et al*. (2016), and Pufall *et al*. (2014), which all found negative impacts of HIV-status on education outcomes that are specific to males. Granted the lack of consensus from extant studies regarding the gendered impact of HIV-related factors on education outcomes, and granted the dearth of such studies in general, there has been little contribution to the literature regarding the potential reasons for a female-specific negative impact of HIV on educational attainment. Regardless, some potential factors are explored here. Firstly, HIV-related stigma, both external and internal, may be more acute for female children and adolescents of school-going age than it is for their male counterparts. A study on HIV stigma among women and adolescent girls in South Africa by Armstrong-Mensah *et al*. (2019) found that women and adolescent girls living with HIV in South Africa experience disproportionate stigma, discrimination, and marginalization as a result of their HIV-positive status. Of particular relevance to this study, the authors find that this stigma extended to gender-compounded discrimination and hostility in educational settings, specifically (Armstrong-Mensah *et al*., 2019). Secondly, familial investment in child health and education may, in the face of HIV diagnosis, be prioritized for male children, given the higher expected returns to investment in education. Armstrong-Mensah *et al*. (2019) note that, in South Africa, women and adolescent girls living with HIV are disproportionately ostracized by their family and peers, specifically due to their reduced eligibility in the marriage market, and their perceived compromised ability to bear children. This reduction in potential marriageability may either result in parents deprioritizing investment in HIV-positive girls' education, or may result in HIV-positive girls dropping out of school in order to seek work granted their familial ostracization. Regardless, this would have a negative impact on educational attainment for HIV-positive female adolescents specifically. Thirdly, the increased likelihood of new HIV infections among adolescent females, relative to adolescent males, may point to a specific impact of *new* HIV infections on educational attainment, relative to the impact of extant infections. As was noted in the summary statistics in Table 1, the mean age of HIV-positive female children and adolescents of school-going age in South Africa (13.65 years) is significantly higher than for their HIV-positive male counterparts (12.31 years), despite their being no statistically significant difference between the mean age of HIV-negative female and male children and adolescents of school-going age. This is likely attributable to a disproportionate number of new HIV infections among older female adolescents, relative to their male counterparts. Indeed, studies have found that earlier sexual debut among South African women is associated with increased HIV seroprevalence and risk of HIV infection (Wand and Ramjee, 2012; Zuma *et al.*, 2010; Harrison *et al.*, 2005), while participation in age-disparate relationships, which typically involve younger female partners with older male partners, are also associated with an increased likelihood, among young adults, of contracting STIs, including HIV (Beauclair *et al*., 2012). In a study of the impact of HIV on South African adolescents' educational attainment, Fotso *et al*. (2018) find that HIV contracted in childhood and adolescence specifically significantly reduced adolescents' educational attainment; thus, the aforementioned over-representation of females among individuals who contracted HIV in adolescence may explain the female-specific nature of the observed impacts of HIV-related factors on educational attainment. Further and more in-depth interrogation of these, and other potential factors underlying the identified gendered impacts of HIV-related factors on education, using qualitative, quantitative or mixed methods research is encouraged.

The final analyses decomposed the identified statistically significant impacts of HIV-related factors on education outcomes separately for female and male children and adolescents of school-going age, specifically. While the male decomposition reported no statistically significant relationships, decomposition of these impacts within female children and adolescents of school-going age specifically suggests that age and being African are the two major contributors to these impacts for HIV-positive girls. The contribution of race to reduced rates of educational attainment may point to an intersectional effect of race and gender on HIV-related contributions to reductions in educational attainment as identified by Sangaramoorthy *et al*. (2017). Alternatively, the contribution of age to reduced rates of educational attainment among females specifically may be attributable to earlier marriage or cohabitation among older adolescent girls, relative to their male counterparts, which may contribute to both lower educational attainment (as married girls are likely to drop out of school) and higher risk of HIV infection (as earlier sexual debut and age-disparate relationships contribute to higher risk of HIV infection). Moreover, this impact may be compounded by lower average age at sexual debut among African adolescent females specifically (Jorgensen-Wells and James, 2020). Once more, further analyses of these potential impacts are encouraged.

This study does have some notable limitations, which must be acknowledged. Firstly, the study uses cross-sectional data from a single iteration of the SABSSM study, since the most recent iteration of the SABSSM survey was the first to assess viral load suppression among those respondents who tested positive, and thus the only iteration in which assessment along the entire HIV treatment cascade could be conducted. Given the cross-sectional nature of the dataset used in this study, any causal inferences made between HIV indicators and educational attainment are naturally weak. Indeed, a study by Durevall, Lindskog and George (2019), which examined the relationship between school attendance and HIV incidence among young women in KwaZulu-Natal using population-based longitudinal data identified the same negative association between education and HIV-incidence as identified here, but subsequently found no causal relationship between the two. Further studies using similar longitudinal data from a cohort study that includes information on HIV indicators and educational attainment – such as the Birth to 30 study, the largest and longest running study of child and adolescent health and development in Africa, could allow for further interrogation of the possible causality between HIV-related factors and educational attainment to be made, and is a recommended avenue for further research.

Secondly, there is a strong likelihood of selection bias in the study, stemming from the high rates of biomedical testing refusal among surveyed respondents – specifically, only 61.1% of eligible respondents provided a blood specimen for HIV-testing. It is likely that this refusal to provide a blood specimen is non-random in nature, with an expectation that individuals who were already aware that they were HIV-positive and individuals that were at higher risk of testing positive for HIV being less likely to provide blood specimens given privacy concerns and stigma-related fears around testing positive for HIV (indeed, these were among the most commonly cited reasons for refusal among those respondents that explained why they refused to participate). This systematic difference between the target population and those included in the study, realized as a non-random exclusion of more at-risk for HIV individuals from the sample, may undermine the validity of some of the conclusions drawn here.

A third limitation arises from the small numbers of observations on treatment which are used in the analyses, particularly further down the HIV treatment cascade. As noted in the summary statistics, the sub-sample of HIV-positive children and adolescents of school-going age was relatively small (551 individuals) with even smaller sub-samples of ART-exposed and virally-suppressed individuals. These small sub-sample sizes may result in models, particularly for the effects of ART exposure and viral load suppression, that have insufficient statistical power to expose relevant effects, undermining the validity of the study.

A fourth limitation is attributable to the potential influence of unobserved indicators which are not captured in the survey data on the outcome variables of interest (i.e. educational attainment). There are various factors which are likely to contribute to educational attainment outcomes for children which are not adequately captured in the data, including parental involvement in education, teacher-student relationships, and adequacy of school facilities. Related to this, a further potential limitation of the study arises from the methodological approach employed, namely propensity score matching. While the approach has been used with increasing frequency in retrospective analyses using observational data such as this one, propensity score matching is limited in its capacity to assess and balance all of the factors that may be relevant to the study, thus resulting in the potential omission of the effects of important factors that may, in turn, affect the outcome of the analyses, generating misleading results (Reiffel, 2020). In light of this, it is recommended that future studies in this area employ more recent and flexible propensity score matching estimators than those employed in this study, thereby allowing for more flexible specification of the covariates to include all interaction. Furthermore, it is recommended that future systematic reviews in this area engage with the differences in methodological approaches *across* studies in the literature, and examine how these differences may explain the differences in the results that have been observed.

Finally, the binary nature of the question used to create the absenteeism dummy variable ("In the last month, have you missed school?") does not allow for detailed interrogation of the expected contribution of HIV-related morbidity to chronic absenteeism, as there is no indication of the total number of school days missed.

8 Conclusion

HIV/AIDS represents the largest contributor to overall disease burden among children and adolescents of school-going age in South Africa. Under the established correlation between health status and education, it is expected that this significant disease burden attributable to HIV/AIDS would manifest as diminished educational attainment outcomes among HIV-affected children and adolescents of school-going age in South Africa. This study has interrogated this expected relationship. The study has found that, among all children and adolescents of school-going age in South Africa, the status of being HIV-positive is significantly associated with a decreased likelihood of being at school, and, among children and adolescents of school-going age that are in school, an increased likelihood of absenteeism and being old for one's grade. Indicators further down the HIV treatment cascade – ART exposure and viral load suppression – do not have an impact on educational attainment, which suggests that the impact of HIV on educational attainment is not associated with HIV-related illness and morbidities, but rather with other HIV-related factors. This study has, further, found that the impacts of HIV-related factors on educational attainment are gendered in nature, with a female-specific negative impact of HIV-status on educational attainment overall, and on school attendance in particular. The fact that HIV-related factors impact on the educational attainment of female children and adolescents of school-going age in South Africa, more so than their male counterparts, provides an important starting point for further research on the gendered nature of the impact of HIV on education in South Africa, and also highlights the need for health and education policies that clearly acknowledge the gendered nature of these impacts and are specifically designed to target HIV-positive adolescent girls and ensure their improved education outcomes.

9 References

Abaasa, Andrew, Yunia Mayanja, Gershim Asiki, Matt A Price, Patricia E Fast, Eugene Ruzagira, Pontiano Kaleebu, and Jim Todd. 2021. “Use of Propensity Score Matching to Create Counterfactual Group to Assess Potential HIV Prevention Interventions.” *Scientific Reports*: 1–11.

Ainsworth M., K. Beegle, and G. Koda. 2005. “The impact of adult mortality and parental deaths on primary schooling in North-western Tanzania.” *Journal of Development Studies* 41: 412–439.

Ainsworth M., and D. Filmer. 2006. “Inequalities in children's schooling: AIDS, orphanhood, poverty, and gender.” *World Development* 34: 1099–1128.

Akbulut-Yuksel, Mevlude, and Belgi Turan. 2010. “Left Behind: Intergenerational Transmission of Human Capital in the Midst of HIV/AIDS.”

Anabwani, Gabriel, Grace Karugaba, and Lesego Gabaitiri. 2016. “Health, Schooling, Needs, Perspectives and Aspirations of HIV Infected and Affected Children in Botswana: A Cross-Sectional Survey.” *BMC Pediatrics* 16 (1): 1–14.

Angrist, Joshua D. 2008. *Mostly Harmless Econometrics: An Empiricist’s Companion*.

Armstrong-Mensah, Elizabeth, Pamela Hernandez, Madina Huka, Angie Suarez, Adeola Akosile, Ann Joseph, and K. Ramsey-White. 2019 "HIV stigma among women and adolescent girls in South Africa: removing barriers to facilitate prevention." *Madridge Journal of AIDS* 3 (1): 69-74.

Bain, Luchuo Engelbert, Clovis Nkoke, and Jean Jacques N. Noubiap. 2017. “UNAIDS 90–90–90 Targets to End the AIDS Epidemic by 2020 Are Not Realistic: Comment on ‘Can the UNAIDS 90–90–90 Target Be Achieved? A Systematic Analysis of National HIV Treatment Cascades.’” *BMJ Global Health* 2 (2): 2016–18.

Bandason, Tsitsi, Lisa F. Langhaug, Memory Makamba, Sue Laver, Karin Hatzold, Stephen Mahere, Shungu Munyati, Stanley Mungofa, Elizabeth L. Corbett, and Rashida A. Ferrand. 2013. “Burden of HIV among Primary School Children and Feasibility of Primary School-Linked HIV Testing in Harare, Zimbabwe: A Mixed Methods Study.” *AIDS Care - Psychological and Socio-Medical Aspects of AIDS/HIV* 25 (12): 1520–26.

Bärnighausen T., V. Hosegood, I. Timaeus, and M. Newell. 2007. “The socioeconomic determinants of HIV incidence: evidence from a longitudinal, population-based study in rural South Africa.” *AIDS* 21 (7): S29–S38.

Beauclair, Roxanne, Reshma Kassanjee, Marleen Temmerman, Alex Welte, and Wim Delva. 2012. "Age-disparate relationships and implications for STI transmission among young adults in Cape Town, South Africa." *The European Journal of Contraception & Reproductive Health Care* 17 (1): 30-39.

Becker, Sascha O., and Andrea Ichino. 2002. “Estimation of Average Treatment Effects Based on Propensity Scores.” *The Stata Journal: Promoting Communications on Statistics and Stata* 2 (4): 358–77.

Bele, Samir D., Sameer Valsangkar, and Trupti N. Bodhare. 2011. “Impairment of Nutritional, Educational Status and Quality of Life among Children Infected with and Belonging to Families Affected by Human Immunodeficiency Virus/Acquired Immune Deficiency Syndrome.” *Vulnerable Children and Youth Studies* 6 (4): 284–92.

Bhargava, Alok. 2005. “AIDS epidemic and the psychological well-being and school participation of Ethiopian orphans.” *Psychology, Health & Medicine* 10 (3): 263-275.

Caliendo, Marco, and Sabine Kopeinig. 2008. “Some Practical Guidance for the Implementation of Propensity Score Matching.” *Journal of Economic Surveys* 22 (1): 31–72.

Case, Anne, and Cally Ardington. 2006. “The impact of parental death on school outcomes: Longitudinal evidence from South Africa.” *Demography* (43): 401-420.

Center for Disease Control (CDC). 2022. “HIV Risk and Prevention.” [Online] Available: https://www.cdc.gov/hiv/risk/.

Champaloux, Steven W., and Deborah R. Young. 2015. “Childhood Chronic Health Conditions and Educational Attainment: A Social Ecological Approach.” *Journal of Adolescent Health* 56 (1): 98–105.

Cluver, L., M. Orkin, M. Boyes, L. Sherr, D. Makasi, and J. Nikelo. 2013. “Pathways from parental AIDS to child psychological, educational and sexual risk: developing an empirically-based interactive theoretical model.” *Social Science & Medicine* 87: 185–193.

Cluver, L., D. Operario, T. Lane, and M. Kganakga. 2012. “‘I Can’t Go to School and Leave Her in So Much Pain’: Educational Shortfalls Among Adolescent ‘Young Carers’ in the South African AIDS Epidemic.” *Journal of Adolescent Research* 27 (5): 581–605.

Cohen, Desmond. 2002. “Human Capital and the HIV Epidemic in Sub-Saharan Africa By.” *ILO Programme on HIV/AIDS*.

Cohen, Joyce, Catherine Reddington, Dawn Jacobs, Regina Meade, Donna Picard, Kathy Singleton, Dorothy Smith, and M Blake Caldwell. 1997. “School-Related Issues Among HIV-Infected Children.” *Pediatrics* 100 (1).

Curley, Jami, Fred Ssewamala, and Chang-Keun Han. 2010. "Assets and educational outcomes: Child Development Accounts (CDAs) for orphaned children in Uganda." *Children and Youth Services Review* 32 (11): 1585-1590.

Currie, Janet. 2009. “Healthy, Wealthy, and Wise: Socioeconomic Status, Poor Health in Childhood, and Human Capital Development.” *Journal of Economic Literature* 47 (1): 87–122.

Dehejia, Rajeev H, and Sadek Wahba. 2002. “Propensity Score-Matching Methods for Nonexperimental Causal Studies.” *The Review of Economics and Statistics* 84 (1): 151–61.

Delva, Wim, An Vercoutere, Catherine Loua, Jonas Lamah, Stijn Vansteelandt, Petra De Koker, and Lieven Annemans. 2009. “Psychological well-being and socio-economic hardship among AIDS orphans and other vulnerable children in Guinea.” *AIDS Care* 21 (12): 1490–1498.

Dlamini, Thembelihle. 2021. “Level of Education and HIV Viral Load Suppression in a Population Under Universal Anti-Retroviral Therapy in Eswatini.”

Durevall, D., A. Lindskog, and G. George. 2019. “Education and HIV incidence among young women in KwaZulu-Natal: An association but no evidence of a causal protective effect.” *PloS One* 14 (3).

Duryea, Suzanne, Sebastian Galiani, Hugo Ñopo, and Claudia Piras. 2007. “The Educational Gender Gap in Latin America and the Caribbean.”’

Ellis, Walter L. 2004. "Factors associated with the academic achievement of perinatally HIV-infected elementary and middle school children." *Negro Educational Review* 55 (1): 51-58.

Fernando, Deepika, Damani De Silva, Richard Carter, Kamini N. Mendis, and Rajitha Wickremasinghe. 2006. “A Randomized, Double-Blind, Placebo-Controlled, Clinical Trial of the Impact of Malaria Prevention on the Educational Attainment of School Children.” *American Journal of Tropical Medicine and Hygiene* 74 (3): 386–93.

Ferreira, Pedro Cavalcanti, Samuel Pessôa, and Marcelo Rodrigues Dos Santos. 2011. “The Impact of Aids on Income and Human Capital.” *Economic Inquiry* 49 (4): 1104–16.

Floyd, S., A.C. Crampin, J.R. Glynn, N. Madise, M. Mwenebabu, S. Mnkhondia, and P.E.M. Fine. 2007. “The social and economic impact of parental HIV on children in northern Malawi: retrospective population-based cohort study.” *AIDS Care* 19 (6): 781-790.

Forston, Jane. 2007. “The Effect of HIV/AIDS on Educational Attainment.” *UNAIDS Economics Reference Group*.

Fortson, Jane G. 2011. “Mortality Risk and Human Capital Investment: The Impact of HIV/AIDS in Sub-Saharan Africa.” *Review of Economics and Statistics* 93 (1): 1–15.

Fortuny, Clàudia, Ángela Deyà-Martínez, Elena Chiappini, Luisa Galli, Maurizio De Martino, and Antoni Noguera-Julian. 2015. "Metabolic and renal adverse effects of antiretroviral therapy in HIV-infected children and adolescents." The Pediatric Infectious Disease Journal 34 (5): 36-43.

Fotso, A. S., O. Banjo, and J. O. Akinyemi. 2018. “HIV and Adolescents’ Educational Attainment in South Africa: Disentangling the Effect of Infection in Children and Household Members.” *SAJCH South African Journal of Child Health* 12 (Special Issue): S4–9.

Gezaghen, Mahlet, and Mukti Upadhyay. 2014. “The Effect of HIV/AIDS on Economic Growth of Southern African Countries.” *Asian Economic and Financial Review* 4 (9): 1146–57.

Institute for Health Metrics and Evaluation (IHME). 2021. “Global Burden of Disease Study Results Tool.” [Online] Available: https://vizhub.healthdata.org/gbd-results/.

Goldberg, Rachel E., and Susan E. Short. 2016. “What Do We Know about Children Living with HIV-Infected or AIDS-Ill Adults in Sub-Saharan Africa? A Systematic Review of the Literature.” *AIDS Care - Psychological and Socio-Medical Aspects of AIDS/HIV* 28: 130–41.

Gordon, Tanja, Frederik Booysen, and Josue Mbonigaba. 2020. “Socio-Economic Inequalities in the Multiple Dimensions of Access to Healthcare: The Case of South Africa.” *BMC Public Health* 20 (1): 1–13.

Gottfried, Michael A. 2014. “Chronic Absenteeism and Its Effects on Students’ Academic and Socioemotional Outcomes.” *Journal of Education for Students Placed at Risk* 19 (2): 53–75.

Grossman, Michael. 2015. “The Relationship Between Health and Schooling: What’s New?” *NBER Working Paper Series*.

Grossman, Michael. 2015. “The Relationship between Health and Schooling.” *IZA Discussion Paper Series*.

Grossman, Michael. 1976. *Household Production and Consumption. The Correlation between-Health and Schooling*. *(Columbia University Press for National Bureau of Economic Research)*. Vol. I.

Guo, Yan, Xiaoming Li, and Lorraine Sherr. 2012. “The Impact of HIV/AIDS on Children’s Educational Outcome: A Critical Review of Global Literature.” *AIDS Care - Psychological and Socio-Medical Aspects of AIDS/HIV* 24 (8): 993–1012.

Gupta, Anil K., Nidhi Rawat, Kuldeep Rai, Surendra Rana, and Sabyasachi Chakraborty. 2013. “Orphan and Vulnerable Children Infected or Affected by HIV/AIDS in Delhi-Situational Analysis and State Government’s Initiative of Household Economic Strengthening.” *Vulnerable Children and Youth Studies* 8 (2): 161–70.

Haas, Steven A., and Nathan Edward Fosse. 2008. “Health and the Educational Attainment of Adolescents: Evidence from the NLSY97.” *Journal of Health and Social Behavior* 49 (2): 178–92.

Habyarimana, James, Bekezela Mbakile, and Cristian Pop-Eleches. 2010. "The impact of HIV/AIDS and ARV treatment on worker absenteeism implications for African firms." *Journal of Human Resources* 45 (4): 809-839.

Hargreaves, James R., and Judith R. Glynn. 2002. “Educational Attainment and HIV-1 Infection in Developing Countries: A Systematic Review.” *Tropical Medicine and International Health* 7 (6): 489–98.

Harrison, Abigail, John Cleland, Eleanor Gouws, and Janet Frohlich. 2005. "Early sexual debut among young men in rural South Africa: heightened vulnerability to sexual risk?." *Sexually Transmitted Infections* 81 (3): 259-261.

Harrison, Sayward E., Xiaoming Li, JiaJia Zhang, Peilian Chi, Junfeng Zhao, and Guoxiang Zhao. 2017. "Improving School Outcomes for Children Affected by Parental HIV/AIDS: Evaluation of the ChildCARE Intervention at 6-, 12-, and 18-Months." *School Psychology International* 38 (3): 264–286.

Heinrich, Carolyn, and Alessandro Maffioli. 2010. "A Primer for Applying Propensity-Score Matching Impact-Evaluation Guidelines." *Impact Evaluation Guidelines*.

Henning, Margaret, Catherine M. Kirk, Emily Franchett, Rose Wilder, Vincent Sezibera, Christian Ukundineza, and Theresa Betancourt. 2018. "Over-Age and Underserved: A Case Control Study of HIV-Affected Children and Education in Rwanda." *Vulnerable Children and Youth Studies* 13 (1): 81–93.

Hensels, I.S., L. Sherr, S. Skeen, A. Macedo, K.J. Roberts, and M. Tomlinson. 2016. "Do not forget the boys – gender differences in children living in high HIV-affected communities in South Africa and Malawi in a longitudinal, community-based study." *AIDS Care* 28 (2): 100–109.

Hill, Andrew, and Andrew Balkin. 2009. "Risk factors for gastrointestinal adverse events in HIV treated and untreated patients." *AIDS Review* 11 (1): 30-38.

Jackson, Margot I. 2009. "Understanding Links between Adolescent Health and Educational Attainment." *Demography* 46 (4): 671–94.

Jockers, Dominik, Sarah Langlotz, Declan French, and Till Bärnighausen. 2021. "HIV treatment and worker absenteeism: Quasi-experimental evidence from a large-scale health program in South Africa." Journal of Health Economics 79.

Jorgensen-Wells, McKell A., and Spencer L. James. 2020. "Social Capital and Age at Sexual Debut: Race Differences in South Africa." Social Sciences 9 (11): 197-212.

Kasirye, Ibrahim, and Eria Hisali. 2010. "The socioeconomic impact of HIV/AIDS on education outcomes in Uganda: School enrolment and the schooling gap in 2002/2003." *International Journal of Educational Development* 30 (1): 12–22.

Kembo, Joshua. 2010. "Social and economic consequences of HIV and AIDS on children: case study of a high-density community in Harare, Zimbabwe." *SAHARA-J: Journal of Social Aspects of HIV/AIDS* 7 (4): 39–46.

Kibret, Solomon. 2018. "The Effect of HIV/AIDS on Household's Healthcare Expenditure and Income in Addis Ababa: A Propensity Score Matching Approach." *HIV/AIDS Review* 17 (1): 103–10.

Kingdon, Geeta. 2002. "The Gender Gap in Educational Attainment in India: How Much Can Be Explained?" *The Journal or Development Studies* 39 (2): 25–53.

Kitara, David Lagoro, Hellen Christine Amongin, Joseph C. Oonyu, and Peter K. Baguma. 2013. "Assertiveness and attitudes of HIV/AIDS orphaned girls towards education in Kampala." *African Journal of Infectious Diseases* 7 (2): 36-43.

Levine, David I, and Gary Painter. 2003. “The Schooling Costs of Teenage Out-of-Wedlock Childbearing: Analysis with a within-School Propensity-Score-Matching Estimator.” *The Review of Economics and Statistics* 85 (4): 884–900.

Marinda, Edmore, Leickness Simbayi, Khangelani Zuma, Nompumelelo Zungu, Sizulu Moyo, Lwando Kondlo, Sean Jooste, et al. 2020. “Towards Achieving the 90–90–90 HIV Targets: Results from the South African 2017 National HIV Survey.” *BMC Public Health* 20 (1): 1–12.

Matshe, I, and O Pimhidzai. 2010. “Macroeconomic Impact of HIV and AIDS on the Zimbabwean Economy: A Human Capital Approach.” *Botswana Journal of Economics* 5 (7).

Mayes, SD, HA Handford, JH Schaefer, CA Scogno, SR Neagley, L Michael-Good and LE Pelco. 1996. “The relationship of HIV status, type of coagulation disorder, and school absenteeism to cognition, educational performance, mood, and behavior of boys with hemophilia.” *The Journal of Genetic Psychology* 157(2): 137-151.

McDonald, Scott, and Jennifer Roberts. 2004. “AIDS and Economic Growth: A Human Capital Approach.”

Meng, Xin, and R. G. Gregory. 2002. “The Impact of Interrupted Education on Subsequent Educational Attainment: A Cost of the Chinese Cultural Revolution.” *Economic Development and Cultural Change* 50 (4): 935–59.

Nichols, Austin. 2008. “Erratum and Discussion of Propensity-Score Reweighting.” *Stata Journal* 8 (4): 532–39.

Nichols, Austin. 2007. “Causal Inference with Observational Data.” *Stata Journal* 7 (4): 507–41.

O’Donnell, Owen, Stephen O’Neill, Tom Van Ourti, and Brendan Walsh. 2016. “Conindex: Estimation of Concentration Indices.” *Stata Journal* 16 (1): 112–38.

O’Donnell, Owen, Eddy van Doorslaer, Adam Wagstaff, and Magnus Lindelow. 2007. *Analyzing Health Equity Using Household Survey Data*. *Analyzing Health Equity Using Household Survey Data*.

Orkin, Mark, Mark E. Boyes, Lucie D. Cluver, and Yuning Zhang. 2014. “Pathways to Poor Educational Outcomes for HIV/AIDS-Affected Youth in South Africa.” *AIDS Care - Psychological and Socio-Medical Aspects of AIDS/HIV* 26 (3): 343–50.

Pan, Wei, and Haiyan Bai. 2018. “Propensity Score Methods for Causal Inference: An Overview.” *Behaviormetrika* 45 (2): 317–34.

Parchure, Ritu, Vijaya Jori, Sanjeevani Kulkarni, and Vinay Kulkarni. 2016. “Educational Outcomes of Family-Based HIV-Infected and Affected Children from Maharashtra, India.” *Vulnerable Children and Youth Studies* 11 (4): 332–38.

Pettifor A., S. McCoy, and N. Padian. 2012. “Paying to prevent HIV infection in young women?” *The Lancet* 379: 1280–1282.

Phillips, Nicole, Taryn Amos, Caroline Kuo, Jacqueline Hoare, Jonathan Ipser, Kevin G.F. Thomas, and Dan J. Stein. 2016. "HIV-associated cognitive impairment in perinatally infected children: a meta-analysis." *Pediatrics* 138 (5).

Phiri, Jane, and John E. Ataguba. 2014. “Inequalities in Public Health Care Delivery in Zambia.” *International Journal for Equity in Health* 13 (1): 1–9.

Poulsen, Helen. 2006. “The Gendered Impact of HIV/AIDS on Education in South Africa and Swaziland: Save the Children’s Experiences.” *Gender and Development* 14 (1): 47–56.

Powers, Daniel A., Hirotoshi Yoshioka, and Myeong-Su Yun. 2011. “mvdcmp: Multivariate decomposition for nonlinear response models.” *The Stata Journal* 11 (4): 556-576.

Pufall, Erica L., Constance Nyamukapa, Jeffrey W. Eaton, Catherine Campbell, Morten Skovdal, Shungu Munyati, Laura Robertson, and Simon Gregson. 2014. “The Impact of HIV on Children’s Education in Eastern Zimbabwe.” *AIDS Care - Psychological and Socio-Medical Aspects of AIDS/HIV* 26 (9): 1136–43.

Ranchhod, Vimal, David Lam, Murray Leibbrandt, and Leticia Marteleto. 2011. “Estimating the Effect of Adolescent Fertility on Educational Attainment in Cape Town Using a Propensity Score Weighted Regression.”

Reiffel, J. 2019. “Propensity Score Matching: The ‘Devil is in the Details’ Where More May Be Hidden than You Know.” *The American Journal of Medicine* 133 (2): 178-181.

Robertshaw, Dianne, and Lee M. Wolfle. 1983. “Discontinuities in Schooling and Educational Attainment.” *Higher Education* 12 (1): 1–18.

Sangaramoorthy, Thurka, Amelia Jamison, and Typhanye Dyer. 2017. "Intersectional stigma among midlife and older Black women living with HIV." *Culture, Health & Sexuality* 19 (12): 1329-1343.

Santelli J., Z. Edelstein, S. Mathur, Y. Wei, W. Zhang, and M. Orr. 2013. “Behavioral, biological, and demographic risk and protective factors for new HIV infections among youth, Rakai, Uganda.” *Journal of Acquired Immune Deficiency Syndrome* 63(3): 393.

Suhrcke, Marc, and Carmen de Paz Nieves. 2010. “The Impact of Health and Health Behaviours on Educational Outcomes in High-Income Countries: A Review of the Evidence.” *International Conference on Education, Social Capital and Health in Oslo 25-26,* no. January 2011: 48.

Souza, Edvaldo, Nicole Santos, Sophia Valentini, Gerlane Silva, and Ana Falbo. 2010. "Long-term follow-up outcomes of perinatally HIV-infected adolescents: infection control but school failure." *Journal of Tropical Pediatrics* 56 (6): 421-426.

Tu X., Y. Lu., X. Li, X. Fang, G. Zhao, X. Lin, and B. Stanton. 2009. “School performance and school behaviour of children affected by acquired immune deficiency syndrome (AIDS) in China.” *Vulnerable Children and Youth Studies* 4 (3): 199–209.

Takuva, Simbarashe, Alison E. Brown, Yogan Pillay, Valerie Delpech, and Adrian J. Puren. 2017. “The Continuum of HIV Care in South Africa: Implications for Achieving the Second and Third UNAIDS 90-90-90 Targets.” *Aids* 31 (4): 545–52.

Wand, Handan, and Gita Ramjee. 2012, "The relationship between age of coital debut and HIV seroprevalence among women in Durban, South Africa: a cohort study." *BMJ Open* 2 (1).

Wilde, Elizabeth Ty, and Robinson Hollister. 2002. “How Close Is Close Enough? Testing Nonexperimental Estimates of Impact against Experimental Estimates of Impact with Education Test Scores as Outcomes.”

World Bank. 2023. “World Development Indicators” [Online] Available: https://databank.worldbank.org/source/world-development-indicators.

Yang, Hwami, Zunyou Wu, Shukai Duan, Z. Li, Xiaoming Li, M. Shen, Ambika Mathur, and Bonita Stanton. 2006. “Living environment and schooling of children with HIV-infected parents in southwest China.” *AIDS Care* 18 (7): 647-655.

Zhou, J., A. Jaquet, E. Bissagnene, B. Musick, K. Wools-Kaloustian, N. Maxwell, A. Boulle, F. Wehbe, D. Masys, J. Iriondo-Perez, and J. Hemingway-Foday. 2012. "Short-term risk of anaemia following initiation of combination antiretroviral treatment in HIV-infected patients in countries in sub-Saharan Africa, Asia-Pacific, and Central and South America." *Journal of the International AIDS Society* 15 (1): 1-12.

Zinyemba, Tatenda, Milena Pavlova, and Wim Groot. 2020. “Effects of HIV/AIDS on Children’s Educational Attainment: A Systematic Literature Review.” *Journal of Economic Surveys* 34 (1): 35–84.

Zinyemba, Tatenda, Milena Pavlova, and Wim Groot. 2021. “Effects of HIV on Gender Gaps in School Attendance of Children in Zimbabwe: A Non-Linear Multivariate Decomposition Analysis.” *Education Economics* 0 (0): 1–19.

Zivin, Joshua Graff, Harsha Thirumurthy, and Markus Goldstein. 2009. “AIDS Treatment and Intrahousehold Resource Allocation: Children’s Nutrition and Schooling in Kenya.” *Journal of Public Economics* 93 (7–8): 1008–15.

Zuma, Khangelani, Geofrey Setswe, Thabile Ketye, Thembile Mzolo, Thomas Rehle, and Ntombizodwa Mbelle. 2010. "Age at sexual debut: a determinant of multiple partnership among South African youth." *African Journal of Reproductive Health* 14 (2): 47-54.