DEPARTMENT OF ECONMICS

LIFE SKILLS EDUCATION: REDUCING SEXUAL RISK BEHAVIOUR AMONG YOUNG WOMEN IN SOUTH AFRICA?

Analysing the effect of life skills education on life skills knowledge, sexual risk behaviour and HIV prevalence

Jaballah Sandra & Wallin Anna

Bachelor's Thesis:	Economics, 15 higher education credits
University:	School of Business, Economics and Law, University of Gothenburg
Program:	Bachelor's Program in Economics
Semester:	Spring semester
Date:	8 June 2020
Supervisor:	Lindskog Annika
Words:	19.222

Abstract

South Africa is considered to be the epicentre of the HIV pandemic and its women are disproportionally affected by the disease. A key strategy to prevent and mitigate the spread of HIV infection is the implementation of life skills education in all primary and secondary schools. The purpose is to increase the knowledge and skills on sexual and reproductive health by providing education, care and support among young people.

This thesis analyses the long-term impact of being exposed to two consecutive life skills education programs, implemented in South Africa between 2000 and 2011, on the level of life skills knowledge, level of sexual risk behaviour and HIV prevalence among young women. The main hypothesis tests whether exposure to the life skills programs decreases the level of sexual risk behaviour through increased level of life skills knowledge. Subsequently, also decreasing HIV prevalence.

The method used is the difference-in-difference, which estimates the effect of the programs across cohorts based on the year of birth and initial level of life skills knowledge across municipalities. The effect of the programs is compared between individuals with little or full exposure to the programs and individuals with no exposure. The findings suggest that the life skills education programs did not have statistically significant effect on the level of life skills knowledge, level of sexual risk behaviour or HIV prevalence. Thus, concluding that the programs have not yielded the desired and anticipated outcomes as specified in this research.

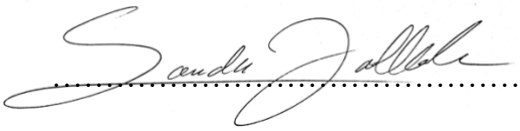
Keywords:

AIDS
Behavioural Health
HIV
Life skills education
Life skills knowledge
Sexual risk behaviour
South Africa

Acknowledgements

We would like to take the opportunity to express our deep gratitude to our supervisor, Annika Lindskog, for her patient guidance, enthusiastic encouragement and useful critiques throughout this research. We would also like to dedicate a special thanks to our families for their support and help in making it possible to finalise this project within the limited time frame. This accomplishment would not have been possible without any of them. Thank you.

…… Gothenburg, 2020-06-08



…… Gothenburg, 2020-06-08

Table of Contents

Table of Figures

Table of Tables

List of abbreviations

2005 LSE program	Life skills education included within the HIV & AIDS and STI National Strategic Plan (NSP) for South Africa 2000-2005
2011 LSE program	Life skills education included within the HIV & AIDS and STI National Strategic Plan (NSP) for South Africa 2007-2011
ART	Antiretroviral therapy
ARV	Antiretroviral drugs
HIV	Human immunodeficiency virus
LSE programs	Joint name for the 2005 LSE program and the 2011 LSE program
LSK	Life skills knowledge
NSP	National Strategic Plan
PrEP	Pre-exposure prophylaxis
SRB	Sexual risk behaviour
STIs	Sexually transmitted infections

1. Introduction

Although the HIV infection rates are decreasing worldwide, the destructive disease remains a serious global issue. In 2018, 1.7 million people were newly infected with HIV (UNAIDS, 2019), out of which 240,000 were residents in South Africa (UNAIDS, 2020). South African women are disproportionally affected by HIV compared to their male counterparts. The disease has proven to disrupt the lives of many women, as they are more likely than men to be discriminated and socially excluded as well as denied employment and other opportunities due to their HIV status (Santos et.al, 2014:10).

The country has for a long time struggled with the pandemic and in 1999, as a response to the peaking HIV infection rates, the South African government in collaboration with several stakeholders developed the five-year *HIV & AIDS and STI National Strategic Plan for South Africa 2000-2005 (NSP)*. One of the main goals of the NSP is to "Promote safe and healthy sexual behaviour" by implementing life skills education (LSE) in all primary and secondary schools (Grade 1 to 12) [hereinafter the 2005 LSE program] (Government of South Africa, 2000:19). In 2006, after the 2005 LSE program had been fully implemented, the Government of South Africa conducted a review on its outcomes. It was concluded that the outcomes had been limited because of difficulties regarding appropriate implementation and the provision of efficient educators' training programs (Government of South Africa, 2006:66; Prinsloo, 2007:159). Based on this review, a new national strategic plan for the period of 2007 to 2011 was drafted [hereinafter the 2011 LSE program]. The new NSP follows the preceding one, with the aim to further develop and strengthen the goals of the previous LSE program[1]. The joint objective of both LSE programs is "to prevent and mitigate the spread of HIV infection, and to provide care and support for learners that are infected and affected by HIV and AIDS" by increasing knowledge, skills and support on sexual and reproductive health (Department of Basic Education, 2019a).

Whether education is an effective method to acquire behavioural change has been a concern for health and development economists as well as policymakers for a long time (Grossman, 2005:1). Knowledge and health are argued to be the two most important sources of capital, and the two are believed to have a mutual impact on each other (Grossman, 2005:12). According to Grossman and Kaestner (1997:74), increased education could have a positive effect on health, however the effect could also be the reversed, where better health results in increased education. This research builds on the belief that increased knowledge through

[1] The 2005 LSE program and the 2011 LSE program will jointly be referred to as "the LSE programs".

education causes a behavioural change, which in turn results in positive health outcomes. More particularly, it argues that life skills education, if targeted and efficient, could constitute an important link between education and the risk of HIV infection. This view is shared by Mabaso (2018:6), who argues that comprehensive sexuality education increases the knowledge of HIV, which in turn reduces the level of sexual risk behaviour and thus, the risk of getting infected by HIV.

In order to assess whether this belief holds in the case of South Africa, this research estimates the effect of the LSE programs on the level of life skills knowledge (LSK), level of sexual risk behaviour (SRB) and HIV prevalence among young women. Previous studies assessing the impacts of the 2005 LSE program in South Africa have found that it increased the knowledge of HIV/AIDS among students. However, the literature seems to disagree on whether there has been an effect on sexual risk behaviours (May et.al. 2004:3; James et.al. 2006:291-292; UNICEF, 2012:84). Moreover, these studies have mainly observed short-to-medium term effects, leaving a gap on the assessment of the long-term impacts. Thus, this research aims to fill the gap by analysing the long-term impacts of the LSE programs on South Africa's women. The analysis is divided into two steps. First, the effect of the programs is compared between individuals with *full* exposure to the programs (they were aged 6 to 9 in 2006), to those with no exposure (they were aged 18 to 21 in 2006). Second, the effect of the programs is compared between individuals with *little* or *full* exposure (they were aged 6 to 17 in 2006), to those with no exposure (they were aged 18 to 29 in 2006). The second step includes 12-year-of-birth dummies, which makes it possible to capture the time dimension of exposure to the programs.

Considering that the LSE programs are full national coverage programs, randomisation is not possible. Consequently, the Difference-in-difference (DiD) estimator is utilised. The identification strategy is based on the fact that exposure to the LSE programs varies by date of birth and the assumption that initial level of life skills knowledge varies between municipalities. Hence, the DiD method is used to estimate the effect of the programs across cohorts and municipalities.

1.1 Aim, research question & hypothesis

This research aims to empirically analyse the long-term impacts of the LSE programs on HIV prevalence among young women, with the main focus on life skills knowledge and sexual risk behaviours. This is a highly interesting subject to study since LSE could constitute an

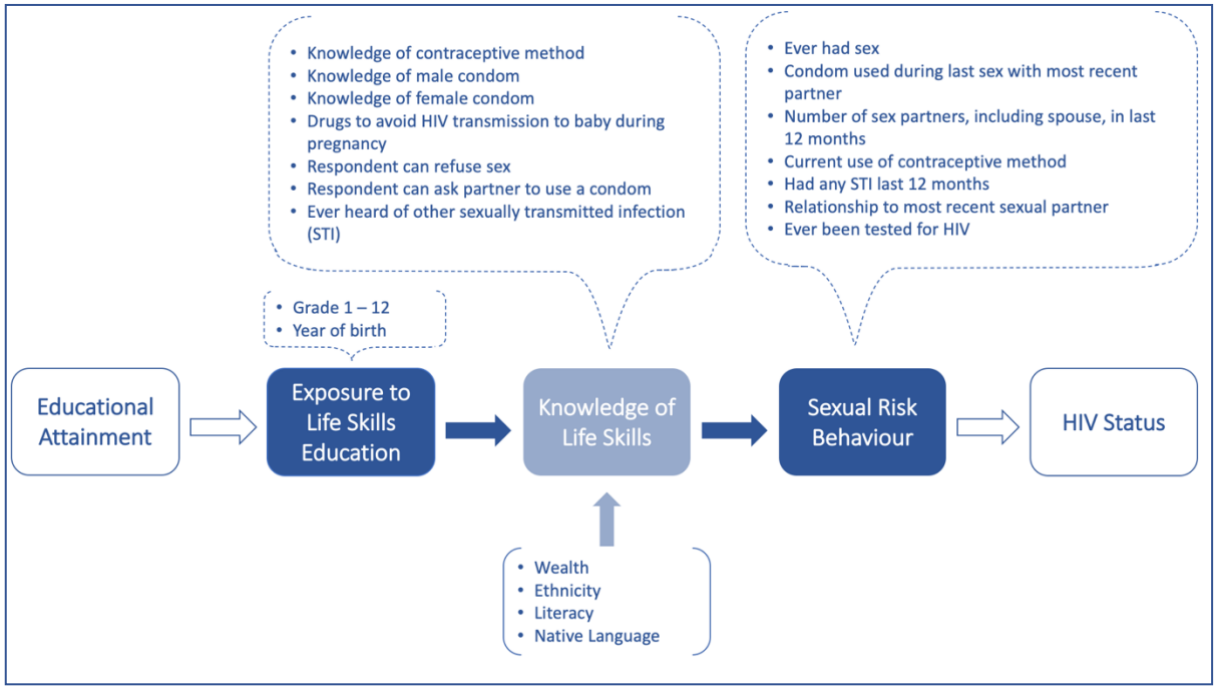
important link between education and the risk of being infected by HIV. The research question and hypothesis are as follows.

RQ*: What are the long-term impacts of exposure to life skills education programs in primary and secondary school on sexual risk behaviour among young women in South Africa?*

H: *Exposure to life skills education decreases an individual's sexual risk behaviour through increased level of life skills knowledge.*

Figure 1 displays the theoretical framework of this study. It provides for a visualisation of the overall relationship, but more importantly an illustration of the main relationship of interest. For clarification purposes, the overall relationship is based on four links; first, individuals who attend school (educational attainment) are exposed to life skills education. Second, individuals who are exposed to life skills education will acquire high knowledge of life skills related issues. Third, individuals who have knowledge of life skills related issues will have low sexual risk behaviour. Finally, individual's with low sexual risk behaviour will have low HIV prevalence.

Figure 1: Theoretical framework



Comment: The main relationship is displayed in the blue boxes. The survey questions within the dashed lines (at the top) displays how each variable is measured. Additionally, four variables controlling for individuals' characteristics are included: wealth, ethnicity, literacy and native language.

1.2 Outline

The outline of this research is as follows. The subsequent section provides for background information on South Africa's education system, the current HIV situation in the country and a discussion on the introduction of life skills education in schools. It also includes a literature review aiming at introducing relevant literature on the theory of education and behavioural health change. The third section describes the data used, which is followed by a discussion on the methodology. The fifth section presents the results of the regression analysis, which are analysed and discussed in the subsequent section. Finally, the research concludes by answering the research question, providing suggestions for future research as well as touching upon some limitations of the study.

2 Background & literature review

This section begins with a background discussion on the education system in South Africa, including challenges and opportunities for teaching and learning. This is followed by general HIV information, an overview of the current HIV situation in the country as well as a discussion of life skills education in South African schools. Finally, the section ends by introducing previous research on the relationship between education and behavioural health change, including a presentation of the identified research gap within the field of research.

2.1 Education in South Africa

At long last, South Africa developed its first curriculum to have full national coverage in 2002 (Department of education, 2002:6). In the following years, the education department developed this further and in 2005, a new curriculum named Curriculum 2005, was implemented at all school levels. The curriculum changed from being subject-based to focus on outcomes-based learning (Moloi & Strauss, 2005:6; Motala & Sayeed Y, 2009). According to the Department of Education (2002:10), "the curriculum attempts to be sensitive to issues of poverty, inequality, race, gender, age, disability, and such challenges as HIV/AIDS". The Curriculum 2005 has been continuously evaluated and further developed; however, its core values and principles have remained.

The education system is divided into four levels: foundation phase (primary education: grade R to 3) [2], intermediate phase (primary education: grade 4 to 6), senior phase (secondary education: grade 7 to 9), and national senior certificate (secondary education: grade 10 to 12) (World Education News Plus Reviews, 2017). Children who normally attend primary and secondary school are between the ages of 5 to 17 (Department of Education, 2008a:2; Statistics South Africa, 2017a:3). However, only grade 1 to 9 are mandatory.

In 2006, the Gross Enrolment Ratio (GER) was 96% for all individuals in the appropriate school-age. For girls, the GER was 95%, compared to 97% for boys (OECD, 2008:20; Department of Education, 2008b). Education has over the past years been remarkably highly prioritized by the government, and the quality of education is considered to be higher compared to other Sub-Saharan African countries. However, due to high levels of school

[2] Grade R is considered as a pre-school grade, but not all schools offer it (World Education News Plus Reviews, 2017).

dropouts and an overall low student performance, the education system remains one of the worst in the world (Kadakia & Macha, 2017).

The average literacy rate in South Africa is slightly lower compared to other developing countries. In 2006, approximately 74.9% of all adults were considered literate, whilst, 10.5% of the adults had no education at all, thus, being completely illiterate. Regardless, the literacy rate was much higher compared to other Sub-Saharan African countries (Department of Education, 2008a:26). Moreover, there is remarkably huge disparities concerning literacy in rural and urban areas. In urban areas, 13% of the sixth graders were reported being functionally illiterate, i.e. could only read a little, whilst the corresponding number was as much as 41% in rural areas. This gap can be attributed to differences in the level of teaching as well as in quality of available resources, such as books, running water and electricity etc. (Kadakia & Macha, 2017; Prinsloo, 2007:158). Furthermore, it is well-known that literacy is fundamental for learning. An illiterate person faces increasingly more barriers to education compared to its peers. Literacy contributes to the development of abilities such as listening, reading, writing as well as critical and creative thinking. All of which help an individual address challenges that may come later in life (ELINET, 2016:3).

South Africa is still one of the world's most unequal countries. In fact, the 90-10 gap was in 2018 the largest one in the world, where as much as 65% of the national income belonged to the richest 10% (Spaull, 2018:1; Statistics South Africa, 2017b:2). Although school enrolment is mandatory and all children have the right to education without discrimination, financially vulnerable individuals are less likely to have the same possibilities to attain their education (Schmitz et.al, 2004:154). In order to mitigate the issue of many students not being able to afford to pay the school fee, the government introduced a “no fee” program that supports low income students with their tuition. Additionally, the “no fee” program also provide for financial support to low income schools to enable proper maintenance and the acquisition of better resources. This initiative has allowed more children to attend school, but unfortunately, the quality of the education in “no fee” schools maintain substantial low, compared to other public and private schools (Motala & Sayeed Y, 2009; Oosthuizen, 2019; U.S. Department of Labor, 2014).

Simultaneously, the high inequality within the country is also reflected between different ethnicities. When the apartheid regime divided people into different racial groups (white, black, Asians with primarily Indian origin and coloured[3]), discriminatory structures impacted

[3] “Coloured” refers to individuals with mixed origin.

peoples' socioeconomic status. Although, the groups are no longer official, the disparities still remain, where non-white individuals are continuously more discriminated and segregated compared to white individuals.

As South Africa consists of many different ethnicities, it is also a multi-language country with as many as eleven official languages. These include Afrikaans and English, as well as nine local languages; Sepedi, Sesotho, Setswana, Siswati, Tshivenda, Xitsonga, isiNdebele, isiXhosa and isiZulu (Department of Education, 2002:19). The three biggest languages spoken at home are isiZulu (25.3%), isiXhosa (14.8%) and Afrikaans (12.2%). Although English constitutes the sixth largest language, with only 8.1% of the population speaking it at home, it together with Afrikaans are the biggest languages taught in school (Government of South Africa, 2020). However, in the 2005 Curriculum, all of the eleven languages are included as teaching languages. Students are expected to learn their home language and at least one additional language. This has not always been the case; the apartheid regime used language in school as another way to segregate the population. Back in 1953, the Bantu Education Department[4] decided that only Afrikaans and English would be the teaching languages in school. Unfortunately, this exclusion still exists, where Afrikaans teaching schools can deny children not speaking Afrikaans their right of being taught in their home language. Nonetheless, most common is that children study their home language until grade 3, but then switch to English or remain with Afrikaans. This seems to affect the learning gap in a negative way since learners with English or Afrikaans as the home language are given an advantage (UNICEF, 2016:94; Department of Education, 2002:19; Spaull 2018:7; Stein N 2017:209).

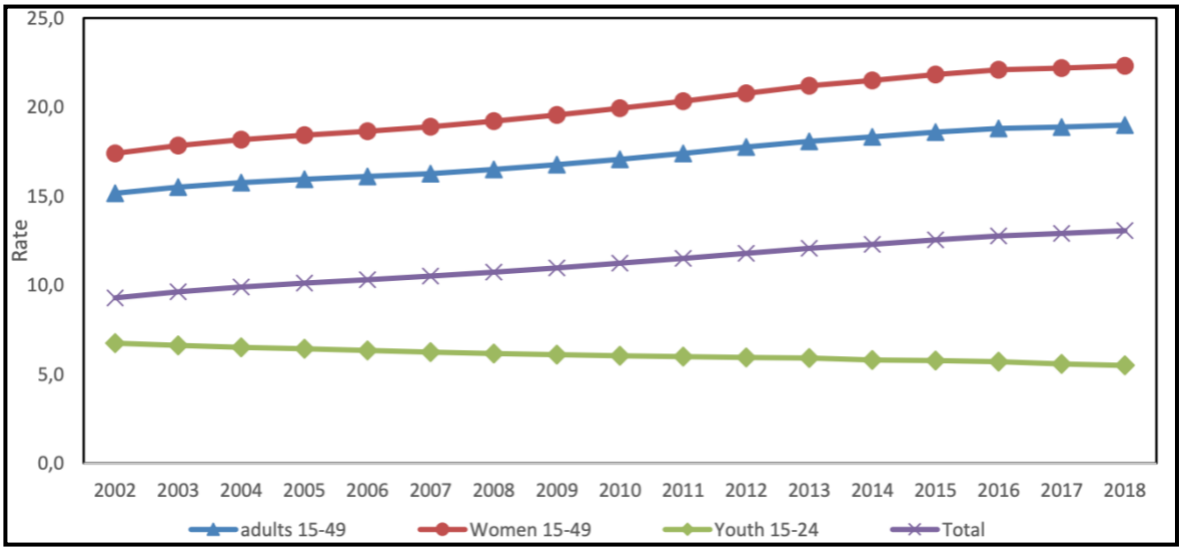
2.2 HIV in South Africa

South Africa has struggled with the HIV/AIDS pandemic ever since the first national case of HIV was reported in 1982 (McNeil, 2019; Simelela et.al, 2015:257). Still today, the disease is rampaging the country. In 2018, South Africa was considered the epicentre of the pandemic with its 7.7 million people living with HIV. Out of these, 63% were adult women and 3% were children (both boys and girls) aged 14 or younger. Women are disproportionally affected by HIV with a prevalence rate of 26% among older women and 11% among younger women. This compared to the prevalence rate among men, where only 15% of older men and

[4] The Bantu Education Department was established in 1953 as a way to uniformly implement discriminatory education structures against black individuals (Kadakia & Macha, 2017).

4% of younger men are living with the disease (UNAIDS, 2020). Figure 2 illustrates the development of the HIV prevalence rate by different groups for the period 2002 to 2018. The number of individuals living with HIV has increased since 2002 in all groups, except for youth aged 15-24 (the green line). However, Zaidi (2013:2304-2305) argues that the overall increase in prevalence can be attributed to the fact that increased access to antiretroviral therapy (ART) has resulted in a substantial decrease in HIV/AIDS mortality rates, rather than an actual increase in the number of newly infected.

Figure 2: HIV prevalence rates, 2002 - 2018



Source: Statistics South Africa (2018:8) Mid-year population estimates 2018, [Online] Statistics South Africa, available at: https://www.statssa.gov.za/publications/P0302/P03022018.pdf, (Accessed: 2020-05-13)

HIV is mainly transmitted through sexual activity or from mother to child during pregnancy, delivery and breastfeeding, as well as through sharing injection equipment (Norrby et.al., n.d). There are several ways to prevent and reduce the risk of being infected by HIV. Besides abstaining from sex, one of the most effective and well-known methods is to engage in safe sexual activity by the use of either male or female condoms. The spread of HIV can also be prevented by careful behaviour, such as using lubricant in addition to condoms during sexual activity, decreasing the number of sexual partners, avoiding sharing injection equipment such as unsterilized syringes, tattoo and piercing equipment as well as by taking pre-exposure prophylaxis (PrEP) if not yet infected or by taking antiretroviral drugs (ARV) if already infected (NSW Health, 2017:1-3). To this day, there does not exist a cure for HIV, however, ART, i.e. the combination of ARV drugs, is identified as an effective method to suppress the virus. According to the World Health Organisation (WHO, 2019), the risk of transmitting the

virus is reduced by 96% if an HIV-infected individual effectively adheres to ART. In 2018, approximately 4.8 million people infected by HIV in South Africa were on ART, out of which, 64% were adult women and 3% were boys and girls aged 14 or younger[5]. Despite large prevention efforts put forward by the government and the community of South Africa, the newly infection rate is still quite large. In 2018, it was estimated that 2.4 thousand individuals were newly infected by HIV, out of which, 58% were adult women and 6% children aged 14 or younger (UNAIDS, 2020).

The Human Sciences Research Council (HSRC, 2019:4-6) has identified several key drivers of the pandemic in South Africa. A few of these include early sexual debut, multiple sex partners (MSP) and age-disparate sexual relationships. According to Richter et.al (2015:304), early sexual debut, i.e. having had first sex before the age of 14, is associated with increased risk of sexual and reproductive health issues, such as HIV and other sexually transmitted infections (STIs). In addition, Wand and Ranjee (2012:5-7) find evidence that the main association between early sexual debut and HIV-infection rates is primarily because of longer duration of sexual life, and thus, the increased likeliness of having had an active sex life with more than one sex partner. Consequently, increasing the risk of having intercourse with a person already infected with HIV.

The proportion of individuals who had an early sexual debut increased from 8.9% in 2002 to 13.6% in 2017. The increase was larger amongst men (6.4%) compared to the increase among women (2.3%) during the same period (HSRC, 2019: xxxviii). Additionally, men are more likely than women to have multiple sex partners (MSP), i.e. more than one sexual partner in 12 months. Nevertheless, in 2017, there was an estimated decrease in the number of men engaged in MSP, however, a steady increase was noted amongst women (HSRC, 2019:142). According to Onoya et.al (2014:104-105), women tend to engage in MSP mainly due to financial and emotional vulnerability and/or pressure from parents to engaging in transactional sex. Similarly, women may also be pressured to engage in age-disparate sexual relationships to meet subsistence needs. Age-disparate relationship refers to a sexual relationship in which there is an age gap of five or more years between the partners (HSRC, 2019: xxxvii). Previous research has found that age-disparate relationships is associated with increased risk of getting infected by HIV (see for instance Evans, et.al, 2016:4). Moreover, women are more likely to be engaged with an older partner compared to men, and it is

[5] Out of all infected adult women (4.7 million), 65% were on ART in 2018. Moreover, out of all children infected (2.6 thousand), 63% were on ART in 2018.

estimated that the tendency of young women to engage in such relationships is not likely to decrease (HSRC, 2019:5).

In general, a person living with HIV and is on effective ART should be able to live a fairly normal and healthy life, not so different from anyone else (Leonard, 2020). However, in South Africa, people living with HIV are often stigmatised and discriminated. Overall, South African women tend to be at higher risk of being infected by HIV compared to their male counterparts. Women are also more often faced with unequal cultural, social and economic opportunities (Mabaso et.al, 2019:2). Subsequently, these gender disparities in addition to the stigma of being infected by HIV have proven to disrupt the lives of many women.

In South Africa, it is quite common that individuals with HIV also find themselves struggling with post-traumatic stress disorder (PTSD) and depression (Spies & Seedat, 2014:6), which in turn have a negative impact on their working and family life, as well as on their access to health clinics and education services (Santos et.al, 2014:7). Santos et.al (2014:6), estimate that approximately 14.4 % of all South African individual's (men and women) infected with HIV abstain from going to clinics, roughly 9.7% quit their jobs and 4.4% withdrew from education services. In addition, more than 11% of the individuals had lost their jobs and 7.7% had been denied an opportunity due to their HIV status. However, only very few individuals had been denied access to education, health services and/or other civil rights. Although Santos et.al's (2014) research analyses both women and men's experience of living with HIV, they stress that the negative experiences are more prominent among women. Females are often blamed for spreading HIV. They are therefore often forced to keep quiet about their HIV status due to fear of violence by their partner. Moreover, HIV-positive women are more likely than men to be discriminated and socially excluded as well as denied employment and other opportunities (Santos et.al, 2014:10).

The South African government has acknowledged the severity of the pandemic and several efforts have been put in place to mitigate the spread of HIV. South Africa is currently working on implementing UNAIDS' 90-90-90 strategy for HIV, which aims at ensuring that 90% of all people with HIV know their status, providing 90% of all HIV infected individuals with ART as well as ensuring that 90% of those that receive ART have virally suppressed HIV (SANAC, 2017:15). Additionally, *South Africa's National Strategic Action Plan for HIV, TB and STIs 2017-2022* [hereinafter NSP 2022] states that HIV should be eliminated as a public health threat by 2030 (SANAC, 2017:3). The NSP 2022 is the country's fourth, most recent and most ambitious plan to combat HIV. It includes several of targeted measures, out of which, many are based on educational efforts that aims at increasing knowledge and

understanding of HIV (SANAC, 2017). This approach has permeated all of South Africa's national strategic plans to combat the disease.

2.3 Introducing life skills in South African schools

In 1999, as a response to the peaking HIV infection rates, the South African government in collaboration with several stakeholders developed the five-year *HIV & AIDS and STI National Strategic Plan for South Africa 2000-2005* (NSP 2005) (Government of South Africa, 2000:5-6). The South African NSP 2005 includes four priority areas; (1) prevention, (2) treatment, care and support, (3) legal and human rights, and (4) monitoring, research and evaluation (Government of South Africa, 2000:5-6). This thesis focuses on the prevention area and goal 1, namely "promote safe and healthy sexual behaviour". In order to reach the objectives of this goal, several strategies were set-up. However, this paper limits itself to the strategy aimed at implementing life skills education (LSE)[6] in all primary and secondary schools (grade 1 to 12) (Government of South Africa, 2000:19). This strategy will hereinafter be referred to as "the 2005 LSE program", with the main objective "to prevent and mitigate the spread of HIV infection, and to provide care and support for learners that are infected and affected by HIV and AIDS" by increasing the knowledge, skills and support on sexual and reproductive health (Department of Basic Education, 2019a).

Having life skills as a subject in South African schools is not new. Already in 1998, the National Coordinating Committee for Life Skills and HIV/AIDS mandated that education of life skills and HIV was to be included in the curriculum, with the main objective to raise gender equality and reduce sexual violence (Adewumi & Adendorff, 2014:460). However, it was concluded that the implementation of the subject was inadequate and inconsistent since each of the nine provinces were individually in charge of the structure and implementation (Magnani et.al, 2005:290). Consequently, the Government of South Africa saw the need of a life skills program with a national coverage. As a result, the 2005 LSE program was created. The 2005 LSE program extends the purpose of life skills by redefining the subject to" the formalized teaching of requisite skills for surviving, living with others and succeeding in a complex society" (May et.al, 2004:8). Moreover, in accordance with the United Nations General Assembly Special Session (UNGASS) on HIV and AIDS, the 2005 LSE program

6 According to UNICEF (2012: Viii), there does not exist a shared definition of "life skills". Nonetheless, it is possible to create a somewhat unified definition based on the most essential features of different actors' attempts to conceptualise the term. Thus, "*Life skills are defined as psychosocial abilities for adaptive and positive behaviour that enable individuals to deal effectively with the demands and challenges of everyday life*" (WHO, 1997:1; UNICEF, 2003).

clearly stresses the importance of targeting youth, since behavioural change is easier to achieve in younger ages and today's youth are crucial for the future of the society, particularly with regards to the economy (Government of South Africa, 2000:25; Government of South Africa, 2006:36). The 2005 LSE program underlines the importance of simultaneously increasing the level of life skills knowledge (LSK) among youth, whilst also changing their sexual behaviours. It is stressed that increased knowledge provided by the life skills education, should lead to a decrease in students' sexual risk behaviours (SRB), and if this cannot be accomplished, the HIV prevalence will remain (Magnani et.al, 2004:297).

Decision-makers from the national, provincial and local district level collaborated to ensure an equal distribution of the resources allocated to the implementation of the 2005 LSE program (Government of South Africa, 2000:29). With these resources, teachers were to be given appropriate training aiming at enhancing their teaching skills on the subject. The Curriculum 2005 specifies that students in grade 1 to 3 should learn about their constitutional rights as well as how to conduct in a responsible and moral way concerning their health and environment. Students are also supposed to obtain basic knowledge about diseases like HIV and AIDS through general health aspects on how to keep your body safe (Department of Basic Education, 2011a). Grade 4 to 6 takes LSE to the next level by including more teaching about the transmission of HIV, how to protect yourself from the disease and how to deal with stigma about HIV and AIDS by addressing common myths. It also addresses issues related to body integrity, safe and unsafe relationships. In secondary school (grade 7-12), education related to sexual behaviour is introduced together with previous HIV and AIDS learning. The level of LSE being taught and level of LSK required steadily increases by each school level, where the more advanced knowledge is taught from grade 4 and upwards (Department of Education, 2002:18: Department of Basic Education, 2011b).

Specifically, students being exposed to the 2005 LSE program are expected to "(1) demonstrate a clear and accurate understanding of sex, sexuality, gender, and STIs, (2) critically identify ways in which HIV/STIs can and cannot be transmitted, (3) identify and evaluate the effectiveness of HIV/STI prevention methods, (4) identify, access, and mobilize sources of assistance within a community, (5) critically evaluate reasons for delaying sexual intercourse or practicing abstinence, (6) respond assertively to pressure for sexual intercourse and unprotected sex, (7) critically evaluate reasons and methods for having protected sex when/if sexually active, (8) accept, cope, and live positively with the knowledge of being HIV positive, (9) show compassion and solidarity towards persons with HIV/AIDS and those affected, (10) provide basic care for people living with HIV and AIDS in the family and

community, and (11) understand and cope with loss and the grieving process." (Magnani et.al, 2004:290; May et.al, 2004:8).

During the five-year implementation period from 2000 to 2005, the rate of newly HIV infections slowed down. However, the prevalence rate was still very high. Although life skills education has been established in all schools since 2005, changing behaviours is not an easy task and the problem remains. Thus, there are still improvements to be made. Previous studies have highlighted that the main shortcoming of the 2005 LSE program can be attributed to difficulties of effectively educating and training teachers on the matter (Government of South Africa, c.2006:66). In an early research conducted on the challenges and achievements of the 2005 LSE program, it was found that teachers tend to be reluctant to discuss life skills related issues with students, since these are considered as sensitive topics (May et.al, 2004:50). In addition, the 2005 LSE program did not fully manage to acknowledge structural differences between schools in urban and rural areas as well as poorer and richer regions, which impeded the implementation (Prinsloo, 2007:164). Nonetheless, an evaluation of the 2005 LSE program made by May et.al (2004:2), confirms that the level of life skills knowledge actually increased when the program was established. The same study also indicates that the increased knowledge had a positive effect on sexual risk behaviour. However, in general, the progress of life skills education is normally difficult to track, and the likelihood of success is very context specific (Aggleton & Clarke, 2012:3; Magnani et.al, 2004:290). Considering the discrepancy among scholars on whether LSE programs generally have the desired effect on sexual risk behaviours, this study argues that the topic remains interesting and highly relevant to study.

In 2006, based on reviews of the implementation and outcomes of the 2005 LSE program, a new national strategic plan for the period of 2007 to 2011 (NSP 2011) was drafted. The new NSP follows the preceding one, with the aim to further develop and strengthen the goals of the previous LSE program[7] (Government of South Africa, c.2006:143; UNICEF, 2012:84). Unlike regular subjects, life skills require a different kind of discipline from teachers. Thus, the main goals of the 2011 LSE program were to maintain and enhance the implementation, but also reinforce educators' teaching abilities on the subject (Government of South Africa, 2006:66; Prinsloo, 2007:159). Although the 2011 LSE program showed some progress, it was affirmed that it did not reach its full potential because of high rates of school dropouts.

[7] The new LSE program included in the *HIV & AIDS and STI Strategic Plan for South Africa 2007-2011*, will hereinafter be referred to as "the 2011 LSE program".

Consequently, the country continued to revise its efforts and a new strategic plan, *the National Strategic Plan on HIV, STIs and TB 2012-2016* (NSP 2016), was developed. The main difference between the NSP 2011 and the NSP 2016, is a shift in focus to instead maintain and improve schooling, i.e. make learners complete grade 12 (Government of South Africa, 2011:17; Kadakia & Macha, 2017). Finally, in 2016, the country developed its most recent and ambitious plan to combat HIV; *South Africa's National Strategic Action Plan HIV, TB and STIs 2017-2022.* Similarly, to its predecessors, the NSP 2022 emphasises the importance of education as a preventive method. However, it aims to scale-up all efforts, including by redefining and extending the subjects of life skills and life orientation in the curriculum[8] (SANAC, 2017:56).

The LSE programs have been argued to be fairly successful, as the they have been given the highest priority and resources to be fully implemented nationwide (Magnani et.al, 2005:290). Out of all LSE programs, the 2005 LSE program and the 2011 LSE program are closest in structure and context as they both focus on effectively implementing LSE in all schools by providing training of educators and targeted teaching. Moreover, since the 2005 LSE program and the 2011 LSE program [hereinafter jointly referred to as "the LSE programs"][9] form the foundation of all LSE programs to follow, this study argues that it is important to assess the effectiveness of the two programs. Thus, the aim of this research is to analyse whether the LSE programs have been effective in increasing LSK, decreasing SRB, and reducing HIV prevalence rates among young South African women.

2.4 Literature review: Education & behavioural health change

Scholars have developed several theories on the relationship between education and health. In 1967, Gary Becker defined a well-developed model of endogenous schooling and health. His model has been repeatedly employed and extended by scholars after him (Duflo, 2001:10). Accordingly, Grossman and Kaestner (1997:74), builds on Becker's model by providing three different theories on the correlation between education and health: (1) there is a causal relationship, where increased education has a positive effect on health, (2) the causal effect is the opposite, meaning that better health tend to increase years of schooling and (3) there is no causal relationship, instead there is a third variable that affect both education and health in the same direction, such as parents' education or wealth.

[8] For more information on the NSP 2022, see the previous section; "HIV in South Africa".
[9] Clarification; only the 2005 and the 2011 LSE programs will jointly be referred to as "the LSE programs".

Throughout the years, different studies have adopted one or more of these three theories of correlation. For instance, Rosenzweig and Schultz (1982:59, cited in Grossman, 2005:12) builds on the third theory by arguing that it is questionable whether education in itself can have an impact on health without being influenced by any other inputs. According to Grossman (2005), this is a good point, however, he claims that knowledge and health are the two most important sources of capitals, where both have an impact on the other. As much as knowledge affect an individual's job career, it also impacts other decisions, such as the use of contraceptive method and demand for medical care. In addition, an individual who has poor health is less likely to educate himself/herself, because of the constraints that comes with having poor health (Grossman, 2005:2). Thus, Grossman (2005:11) suggests that "an increase in knowledge capital or schooling raises the efficiency of the production process in the nonmarket or household sector, just as an increase in technology raises the efficiency of the production process in the market sector".

Although Stacey (1998:56) agrees with Grossman and Kaestner (1997) regarding the importance of analysing education beyond its economic impact, she remains hesitant on whether education through governmental intervention always have the desired effect on behavioural changes. Thus, she argues that public policy interventions aimed at changing behaviours are often costly and the success rates are seldom very high due to factors that are unaccounted for, such as self-discipline and problem solving etc. (Stacey, 1998:60). In contrast, Mabaso (2018:6), who based on a national survey conducted in South Africa, finds that sexual education tends to increase HIV knowledge, which in turn decreases an individual's sexual risk behaviour, and therefore, also reduces the risk of attracting HIV.

However, acquiring behavioural change is not an easy task. The neoclassical model of economics assumes that fully informed individuals act rationally and according to their own self-interest. The benefits of investing in preventive measures against negative health outcomes should therefore outweigh the costs of that investment (Dupas, 2011:428-429). In her research Dupas (2011:428-430), considers a model in which individuals have to choose between either investing in preventive or remedial health care. The former reduces the risk of bad health shocks, whilst the latter restore the health stock when poor health conditions already exist. The decision of which health care to invest in depends on the individual's level of information. Whether an individual invests in preventive care will depend on how effective the individual believes the investment will be in reducing negative shocks. The decision to invest in remedial care will instead be based on the individual's belief of what sickness they face and how accessible they believe effective remedial care is.

According to Dupas (2011:428), the behaviours of individuals in developing countries depart from the traditional neoclassical model. Based on her model; imperfect information could result in the individual over- or underinvesting in preventative and remedial care. Therefore, she argues that imperfect information on illness prevention or on the effectiveness and cost-effectiveness of preventative behaviours could constitute the main reason for why individuals in developing countries tend to underinvest in preventative health care (Dupas, 2011:430-431).

Although information could be an effective method to acquire behavioural health change, the type of information and by whom it is provided seems to also matter. In her research, Dupas (2011:431-433) provides several empirical evidences of how and why the source of information are of importance. For instance, she refers to a study conducted by herself on the effect of a relative-risk information campaigned aimed at changing risk behaviours among teenagers in Kenya. Her results suggest that it is difficult to affect behaviours with measures that only incorporate one type of preventative behaviour. Instead, she argues that in order to effectively impact behaviours, it is necessary to provide comprehensive risk and prevention information. Dupas (2011:433) also presents the example of India's unsuccessful attempt at increasing oral rehydration therapy among children in the early 1990s. In this example, she argues that the campaign was unsuccessful because of previously bad track records of the government, which caused the population to distrust government initiatives (Dupas, 2011:433).

Additionally, it is also important that the information target the right group. In his research, de Walque (2007) studies the effect of an HIV/AIDS information campaign on individuals with different levels of educational attainment in Uganda. His results suggest that more educated people are more susceptible to information compared to those with less education. More specifically, he finds evidence of a positive association between education and change in sexual risk behaviour. Although education is not completely exogenous, he concludes that increased education tends to increase the use of condoms and thus, decrease HIV prevalence among young individuals (de Walque, 2007:712-713).

The main takeaways from the brief overview presented above is that health behaviours of individual's in developing countries could be responsive to information, if it is comprehensive, provided by a reliable source and communicated to the right target group. Thus, this research builds on the literature arguing that increased knowledge through education causes a behavioural change, which in turn results in positive health outcomes. Subsequently, it hypothesises that increased life skills education increases the knowledge

about sexual and reproductive health related issues, which leads to a decrease in sexual risk behaviour. Consequently, reducing the risk of being infected by HIV. In order to test this hypothesis, we aim to estimate the effect of being exposed to South Africa's LSE programs on the level of LSK, level of SRB and HIV prevalence. Previous studies assessing the impacts of the 2005 LSE program have found that it increased the knowledge of HIV/AIDS among students. However, the literature seems to disagree on whether there has been an effect on condom use, or on any other sexual risk behaviours. Moreover, these studies have mainly observed short-to-medium term effects (May et.al. 2004:3; James et.al. 2006:291-292; UNICEF, 2012:84), leaving a gap on the assessment of the long-term impacts. Thus, this research aims to fill the gap by analysing the long-term impacts of the LSE programs on HIV prevalence among young women, with the main focus on life skills knowledge and sexual risk behaviours.

3 Data

The purpose of this section is to provide an overview of the datasets and variables used to reach the aim and answer the research question. To get a clear understanding of the variables, descriptive statistics are also presented.

3.1 Dataset

In order to estimate the long-term effects of the LSE programs, we use the latest available data containing information on educational attainment and HIV, produced in 2016 and 2017. The main data source used in this study is the South Africa Demographic and Health Survey (SADHS 2016) conducted in 2016. It is nationally representative since it contains key indicators for the country as a whole and for all of South Africa's nine administrative provinces and municipalities, collected from a sampled population where certain clusters are over-sampled whilst others are under-sampled. The data includes information on individual's socioeconomic status, educational attainment, knowledge of HIV and current HIV status as well as on awareness and use of contraceptives, and sexual behaviour. The main units of analysis are women and the total sample size is 8,737 women aged 15-49 (National Department of Health et.al, 2019:1-3).

The second dataset used in this study is the SADHS conducted in 1998 (SADHS 1998). Similarly, to the SADHS 2016, the SADHS 1998 is nationally representative. It includes information on individual's socioeconomic status, their educational attainment, knowledge of HIV and on awareness and use of contraceptives, and sexual behaviour. The main units of analysis are women and the sample size is 11,735 women aged 15-49 (National Department of Health/ Statistics South Africa and Macro International, 2000: 4-5).

3.2 Variables

3.2.1 Dependent variables

As this study has hypothesised a stepwise impact of the LSE programs, three different relationships are analysed. Thus, three different dependent variables are utilised, and these are described below.

3.2.1.1 2016 LSE index

The first dependent variable is an index measuring the level of life skills knowledge in 2016 and will be referred to as the *2016 LSK index*. It was created by using several relevant variables that individually captures at least one of the eleven aspired outcomes of the 2005 LSE program[10]. However, jointly they capture the most important aspects of the LSE program. The selected variables are presented in table 1 including a description of how they are coded[11]. All variables are coded so that a high value corresponds to high level of LSK, whilst a low value indicates low level of LSK. Subsequently, the index is measured in the same manner. Finally, in order to ensure fairly normally distributed residuals, the *2016 LSK index* is logged in all regressions.

Table 1: Variables constituting the 2016 LSK index

Variable	Coding
Knowledge of contraceptive method	1 = Do not know about contraceptive method 2 = Have knowledge about contraceptive method
Knowledge of male condom	1 = Do not have knowledge about male condoms 2 = Have knowledge about male condoms
Knowledge of female condom	1 = Do not have knowledge about female condoms 2 = Have knowledge about female condoms
Drugs to avoid HIV transmission to baby during pregnancy	1 = Do not have knowledge of how HIV is transmitted 2 = No, HIV cannot be transmitted during pregnancy 3 = Yes, HIV can be transmitted during pregnancy
Respondent can refuse sex	1 = Do not have a regular partner 2 = No, cannot refuse sex 3 = Yes, can refuse sex
Respondent can ask partner to use a condom	1 = Do not have a regular partner 2 = No, cannot ask partner to use condom 3 = Yes, can ask partner to use condom
Ever heard of sexually transmitted infection (STI)	1 = No, never heard of STIs 2 = Yes, have heard of STIs

3.2.1.2 2016 SRB index

The second dependent variable is an index measuring the level of sexual risk behaviour [hereinafter *2016 SRB index*]. The index is based on seven variables that each capture different aspects of sexual risk behaviour. These are presented in table 2, including a description of how they are coded[12]. Noteworthy, is that the variables are coded so that a high value indicates low SRB, whilst a low value indicates high SRB. Thus, the index is coded in the same way. Lastly, the *2016 SRB index* is also logged to ensure a more normalised distribution of the residuals in all regressions.

[10] See the background section for the eleven aspired outcomes of the 2005 LSE program.
[11] Descriptive statistics of each variable is presented in panel A in table A4 in appendix.
[12] Descriptive statistics of the variables are included in panel B in table A4 in appendix.

Table 2: Variables constituting the 2016 SRB index

Variable	Coding
Current use of contraceptive method	1 = Do no use any contraceptive method 2 = Uses contraceptive method
Condom used during last sex with most recent partner	1 = No, condom was not used during last sex 2 = Yes, condom was used during last sex 3 = Is not sexually active
Ever had sex	1 = Yes, have had sex 2 = No, have never had sex
Number of sex partners, including spouse, in last 12 months	1 = Have had more than one sex partner 3 = Have not had a single sex partner or only one
Had STI in last 12 months	1 = Yes, have had STI or do not know 2 = No, have not had STI
Ever tested HIV	1 = No, have never been tested for HIV 2 = Yes, have been tested for HIV
Relationship to most recent sexual partner	1 = It was not a regular partner 2 = It was a regular partner 3 = Is not sexually active

3.2.1.3 HIV prevalence

The third and final dependent variable *HIV* measures HIV prevalence rates. It is based on a dichotomous variable estimating number of individuals living with HIV, where a value of 0 indicates HIV positive, whilst a value of 1 indicates HIV negative. Thus, the variable is coded in the same direction as both the *2016 LSK index* and the *2016 SRB index*. Moreover, since the HIV variable is dichotomous, the fitted model is a linear probability DiD model where the estimated effect is interpreted as the change in the probability when the dependent variable equals to 1, whilst holding all other variables constant (Hippel, 2015).

3.2.2 Main independent variables

3.2.2.1 Cohort of exposure

The main focus lies on young women who have been partly or fully exposed to the LSE programs. Since the 2005 LSE program had its implementation period between 2000 to 2005, it is safe to assume that the program was fully implemented in all schools by 2006. However, in 2007, the 2005 LSE program was revised into the 2011 LSE program, which was to be fully implemented by 2011 (Government of South Africa, 2000:5-6). As has been discussed in the background section, there are no noteworthy changes between the two programs. Therefore, the change from the 2005 program to the 2011 program does not impose any significant constraints on our model.

An individual's year of birth determines her exposure to the LSE programs. Individuals aged 6 to 17 in 2006 have had little or full exposure, thus, these are considered as the cohort of exposure. However, only individuals aged 6 to 9 in 2006 (they were 16 to 19 in 2016 when the data was collected) have had the possibility to have been fully exposed[13]. Therefore, a dummy variable called the *young cohort of exposure* is created, where a value of 1 represents individuals who had *full* exposure to the LSE programs and a value of 0 represents the control cohort, i.e. those with no exposure to the programs (they were 18 to 21 in 2006 and 28 to 31 in 2016). Additionally, by comparing between individuals with *little* or *full* exposure to the programs to those with no exposure (they were 18 to 29 in 2006) using 12-year-of-birth-dummies, it is also possible to capture the time dimension of exposure.

Another limitation is made on the basis of each individual's level of educational attainment. Both LSE programs are designed to be implemented from grade 1 to grade 12, however, school enrolment is only mandatory until grade 9. Therefore, this study only includes individuals that have finished at least grade 9, and the cohort of exposure is limited to individuals that have had at least 1 year of exposure to the programs. Thus, accounting for a large proportion of school dropouts. At most, 4,648 females are included in the regressions. Table 3 below presents how many years an individual has been exposed to the LSE programs in 2016. The youngest ones (6-years-olds) have in 2016 had 9 to 12 years of exposure to the LSE programs, whilst the oldest ones (17-year-olds) have had 1 year of exposure to the LSE programs.

[13] The more advanced LSE is taught from grade 4 and upwards, thus, by being fully exposed to the programs means that the individual has studied until at least grade 9 and has had the possibility to have started or finished grade 12 by 2016.

Table 3: Individuals' exposure to the LSE programs by age

	Grade in 2006	Years exposed to the LSE programs in 2016
Age in 2006		
6	1	9-12
7	2	9-11
8	3	9-10
9	4	9
10	5	8
11	6	7
12	7	6
13	8	5
14	9	4
15	10	1-3
16	11	1-2
17	12	1

Comment: The table presents individuals age and the grade they started in 2006, as well as, how many years they should have been exposed to the LSE programs in 2016.

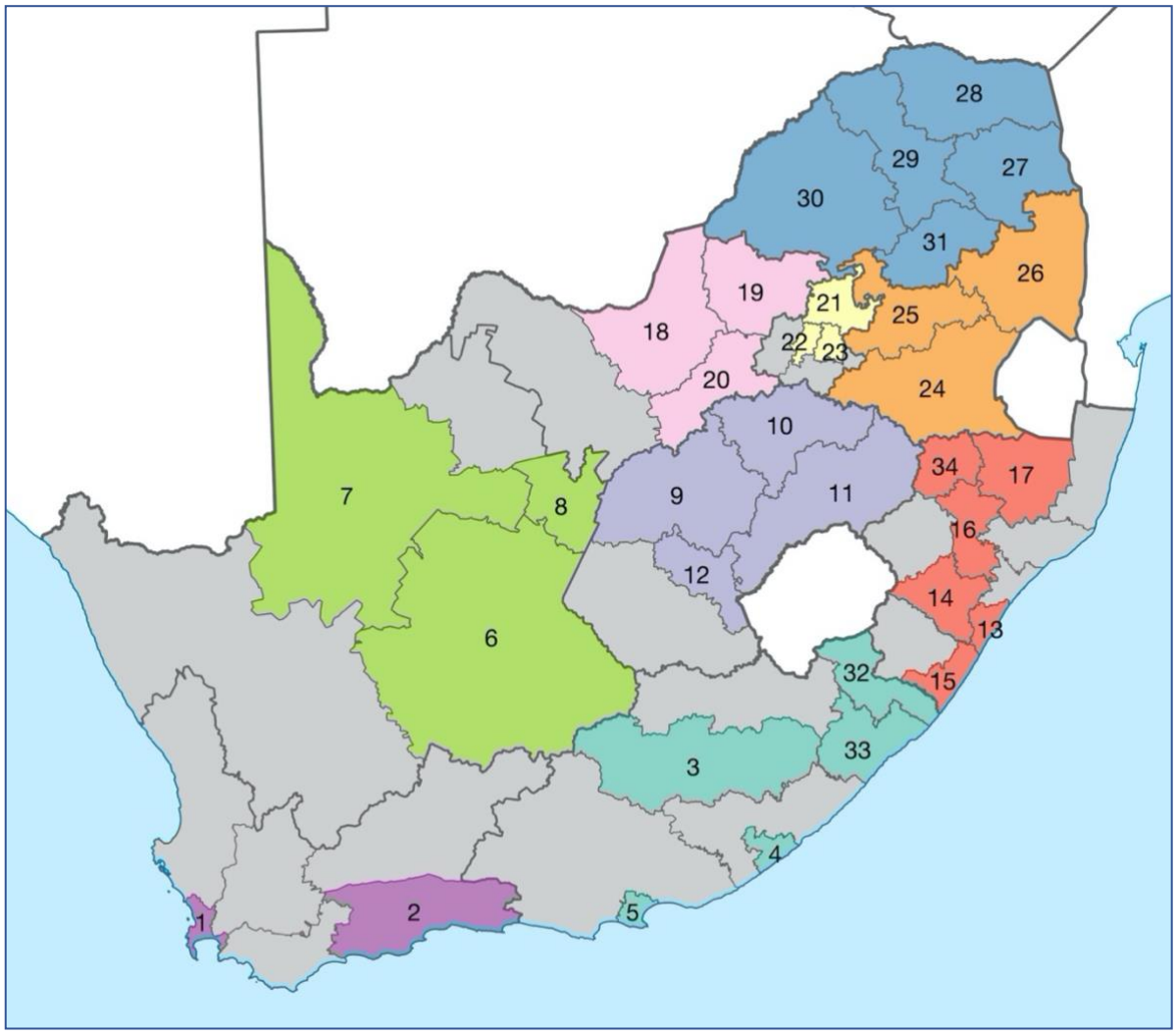
3.2.2.2 Municipalities

This research also considers variation in initial life skills knowledge across municipalities[14]. The selection of the included municipalities is done on the basis of availability and number of observations in the SADHS 2016. For the purpose of not having too few observations, a limit is set to 25 per municipality. Thus, municipalities with less than 25 observations are excluded. In total, 34 out of 52 municipalities are included and these have been coded with a number between 1 to 34. Figure 3 displays the municipalities that are included in the analysis, where each municipality has been assigned a colour. Municipalities with the same colour belongs to the same province[15] (see also table A1 in appendix for a clearer overview of the municipalities, including a presentation of the numerical coding).

[14] Municipalities constitute the second and third governmental layer, following provinces. There are in total eight metropolitan municipalities and 44 district municipalities. The main difference between these two is that metropolitan municipalities have exclusive decision-making power over its local governments, whilst district municipalities share the executive power with its local governments (Matebesi, 2017:10).

[15] For example, municipality 1 is City of Cape Town and municipality 2 is Garden Route District Municipality. Both of which have the colour mauve which indicates that they belong to the Western Cape Province.

Figure 3: Municipalities included in the analysis



Comment: Municipalities included in the analysis are colour coded where municipalities with same colours belongs to the same province. Each municipality is given a number between 1-34. The number coding is presented in table A1 in appendix.

The first difference is enabled by the division of municipalities based on the level of life skills knowledge in 1998. This division is done using an index based on variables capturing life skills related knowledge of females aged 18 to 30[16] in 1998 from the SADHS 1998 [the index is hereinafter referred to as the *1998 initial LSK index*]. Table 4 below presents the variables that constitute the *1998 initial LSK index* and a description of how each variable is coded (see also table A3 in appendix for descriptive statistics of the variables). The selection of the variables to the *1998 initial LSK index* was based on the eleven aspired outcomes of the 2005 LSE program[17]. As can be noted from the table below, each variable is coded so that

[16] Women aged 18-30 are chosen, because they incorporate the main age spectrum of the cohort of exposure in 2016.

[17] See the background section for the eleven aspired outcomes of the 2005 LSE program.

a high value corresponds to high level of LSK, whilst a low value indicates low level of LSK. Accordingly, the index is coded in the same way.

Table 4: Variables constituting the 1998 initial LSK index

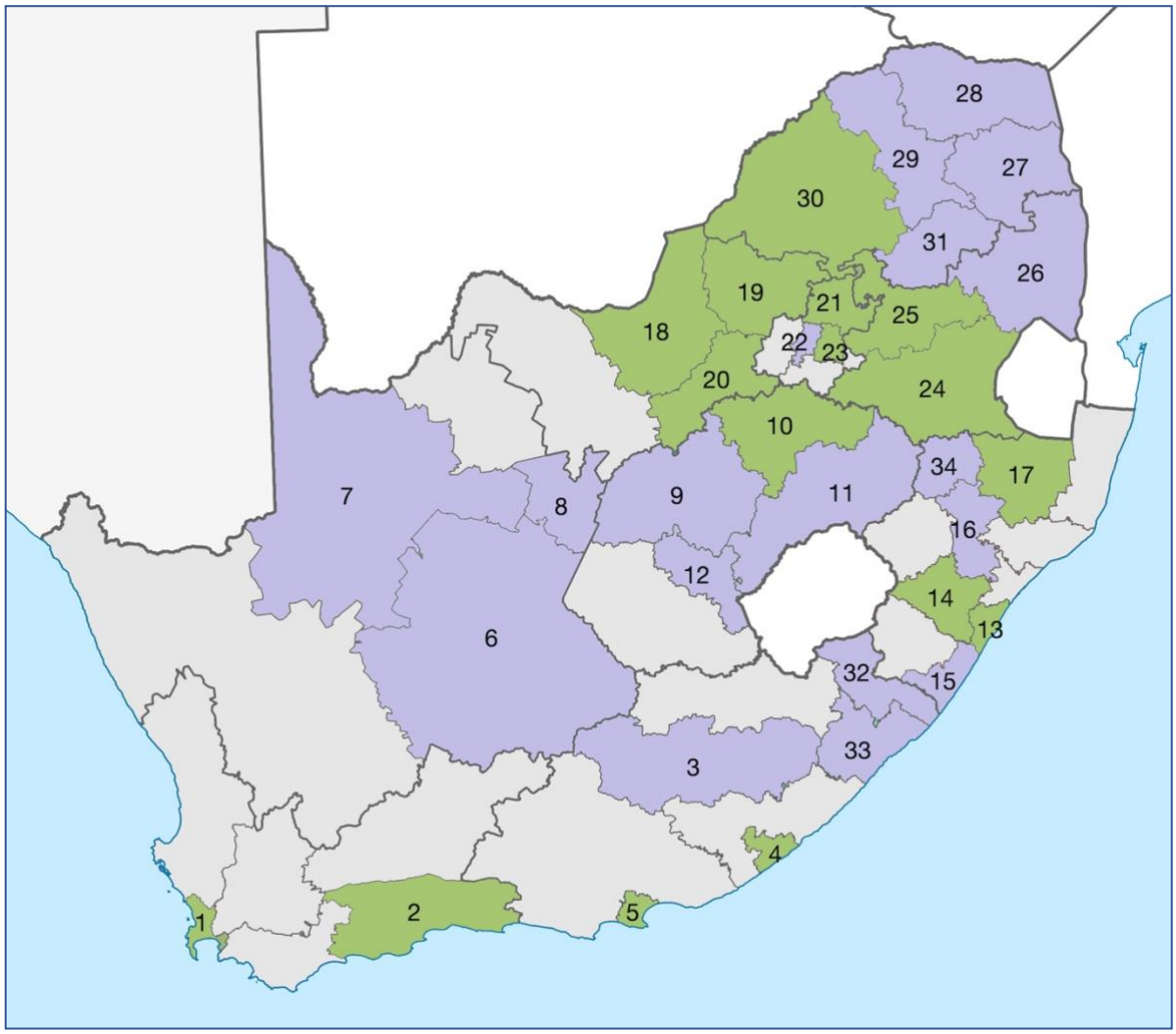
Variable	Coding
Knowledge of contraceptive method	1 = Do not know about contraceptive method 2 = Have knowledge about contraceptive method
Knowledge of condom	1 = Do not have knowledge about condoms 2 = Have knowledge about condoms
Source of condom	1 = Gets condom from illegitimate source 2 = Gets condom from legitimate source 3 = Do not use condom, because not sexually active
Don't use condom because it's a waste of sperm	1 = Yes, it is a waste of sperm 2 = No, it is not a waste of sperm 3 = User of condom
Don't use condom because don't know about condom	1 = Do not have knowledge about condoms 2 = Have knowledge about condoms 3 = User of condom
Don't use condom because don't know how to use	1 = Do not know how to use condom 2 = Have knowledge of how to use condoms 3 = User of condom
Don't use condom because don't know any source	1 = Do not know where to get condoms 2 = Have knowledge of where to get condoms 3 = User of condoms
Don't use condom because embarrassed to get	1 = Yes, it is embarrassing to get condoms 2 = No, it is not embarrassing 3 = User of condoms
Don't use condom because it will be lost inside	1 = Yes, it is a risk of losing the condom inside 2 = No, there is not a risk of losing the condom inside 3 = User of condoms
Don't use condom because low risk of STD	1 = Yes, there is a low risk of STDs 2 = No, there is not a low risk of STDs 3 = User of condom
Don't use condom because it is not cool/trendy	1 = Yes, condoms are not cool/trendy 2 = No, condoms are cool/trendy 3 = User of condoms
Don't use condom because partner dislikes	1 = Yes, partner dislike the use of condoms 2 = No, partner do not dislike the use of condoms 3 = User of condoms
Don't use condom because religion prohibits	1 = Yes, religion prohibits the use of condoms 2 = No, religion does not prohibit the use of condoms 3 = user of condoms
Protect from HIV by avoid mosquitos	0 = Have never heard of HIV 1 = Yes, mosquitos can transmit HIV 2 = No, mosquitos cannot transmit HIV
Protect from HIV by having safe sex	0 = Have never heard of HIV 1 = No, safe sex is not a protective method 2 = Yes, safe sex is a protective method
Protect from HIV by using condom during sex	0 = Have never heard of HIV 1 = No, condom is not a protective method 2 = Yes, condom is a protective method
Protect from HIV by using clean needles during injections	0 = Have never heard of HIV 1 = No, it is not necessary to use clean needles 2 = Yes, it is necessary to use clean needles
Protect from HIV by avoid touching a person with AIDS	0 = Have never heard of HIV 1 = Yes, HIV can be transmitted by touching a person with AIDS 2 = No, HIV is not transmitted by touching a person with AIDS
Protect from HIV by having a good diet	0 = Have never heard of HIV 1 = Yes, a good diet protects from HIV 2 = No, a good diet does not protect from HIV
Protect from HIV by avoiding public toilets	0 = Have never heard of HIV 1 = Yes, HIV can be transmitted by using public toilets

	2 = No, HIV cannot be transmitted by using public toilets
Protect from HIV by avoiding sharing food with person with HIV	0 = Have never heard of HIV 1 = Yes, HIV can be transmitted by sharing food with a person with AIDS 2 = No, HIV cannot be transmitted by sharing food with a person with AIDS
Protect from HIV by avoiding sharing razor blades	0 = Have never heard of HIV 1 = No, it is ok to share razor blades 2 = Yes, do not share razor blades
Knowledge of how to protect from HIV	0 = Have never heard of HIV 1 = No, do not have knowledge of HIV preventive methods 2 = Yes, have knowledge of HIV preventive methods
A healthy-looking person can have HIV	0 = Have never heard of HIV 1 = No, a healthy-looking person cannot have HIV 2 = Yes, a healthy-looking person can have HIV

After using the *1998 initial LSK index* to rank the municipalities from high to low life skills knowledge, the median is used as the cut-off for the division into "high initial LSK" and "low initial LSK" municipalities[18]. Figure 4 illustrates this division, where green colour indicates "high initial LSK" municipality, whereas purple indicate "low initial LSK" municipality. In total, 16 municipalities were classified as "high initial LSK" municipalities, whilst 18 municipalities were classified as "low initial LSK municipalities. A systematic trend can be noted by comparing figure 3 and figure 4. As can be seen, municipalities that belong to the same province tend to also have similar levels of initial LSK. For example, municipality 1 is City of Cape Town and municipality 2 is Garden Route District Municipality. Both of which belong to the Western Cape province and are classified as "high initial LSK" municipalities. Finally, municipalities are included in the models as a dummy variable called *LSK intensity,* where a value of 1 represents "high initial LSK" municipalities, whilst a value of 0 signifies "low initial LSK" municipalities.

[18] The "high initial LSK" municipalities have high initial life skills knowledge compared to municipalities that are considered as "low initial LSK" municipalities.

Figure 4: "High initial LSK" and "low initial LSK" municipalities in 1998



Comment: Municipalities included in the analysis are colour coded where "high initial LSK" municipalities are given the colour green and "low initial LSK" municipalities are given the colour purple. Each municipality is coded with a number between 1-34. The coding is presented in table A2 in appendix.

3.2.2.3 Interaction effect

This study includes several interaction terms, hence, allowing for variation of the marginal effect between the variables constituting the interaction term and thus, the possibility to estimate the average effect of the programs (Aneshensel, 2013:320–321). First, we include an interaction between the variable *young cohort of exposure* and *initial LSK intensity.* Thereby, making it possible to estimate the effect of the programs on the young cohort of exposure as a whole. Thereafter, we replace the interaction between the variable *young cohort of exposure* and *initial LSK intensity* with 12 interaction terms created by multiplying 12 different age dummies, ranging from age 6 to age 17, with *initial LSK intensity.* Thus, making it possible to also distinguish the effect of the programs on each individual age group.

3.2.3 Control variables

This thesis also includes several control variables that may impact an individual's ability to learn and develop comprehensive knowledge of life skills. The first included control is *literacy* which is coded on a three-categorical scale, where a value of 1 indicates that the individual cannot read, a value of 2 means that the individual can only read a little, whilst a value of 3 indicates that the individual can read. However, in the fitted models, value 1, i.e. cannot read, is used as the baseline and, thus, it is omitted.

Moreover, to account for the possible effects that the level of income can have on an individual's ability to receive proper education, this study also includes *wealth* as a control variable[19]. The variable *wealth* is coded in a similar manner as *literacy,* i.e. on a three-categorical scale, where a value of 1 is equal to low income, a value of 2 corresponds to middle income and finally, a value of 3 indicates high income. The first category, i.e. low income, is omitted from the model since it is used as the baseline category.

The third control variable included is *language* which is estimated on a three-categorical scale, where the first categorical level corresponds to the individual having English as first language. The second category indicates that the individual has Afrikaans as home language, whilst the third category indicates that the individual has a local language as native tongue. The third category is used as the baseline; thus, it is omitted from the models.

Finally, we also control for *ethnicity* which is the only variable on a four-categorical scale. The categories are as follows: 1 indicates that the individual is black/African, 2 corresponds to the individual being white, 3 denotes that the individual is coloured, and 4 corresponds to the individual being Indian/Asian[20]. The first category, i.e. "white", is omitted from the model, since it is used as the baseline group.

[19] It is possible that a person's wealth changes across time and place. However, this is difficult to account for due to data restriction. Nevertheless, it is considered during the analysis of the result.
[20] The categories are directly derived from the original variable included in the (SADHS 2016). Thus, no alterations have been made to the labels.

3.2.4 Descriptive statistics

Table 5, panel A to D provides for descriptive statistics of the variables included in this study.

Table 5: Descriptive Statistics

	Observations	Mean	Standard deviation	Minimum	Maximum
Panel A: Individual level					
Education in years (whole sample)	6,842	11.292	1.625	9	16
Sample of interest$_1$	5,027	0.494	0.500	0	1
Education in years (sample of interest$_1$)	5,027	11.310	1.561	9	16
Individuals aged 6 in 2006	138 (5,027)	6	0	6	6
Individuals aged 7 in 2006	214 (5,027)	7	0	7	7
Individuals aged 8 in 2006	198 (5,027)	8	0	8	8
Individuals aged 9 in 2006	219 (5,027)	9	0	9	9
Individuals aged 10 in 2006	254 (5,027)	10	0	10	10
Individuals aged 11 in 2006	206 (5,027)	11	0	11	11
Individuals age 12 in 2006	244 (5,027)	12	0	12	12
Individuals age 13 in 2006	217 (5,027)	13	0	13	13
Individuals age 14 in 2006	258 (5,027)	14	0	14	14
Individuals age 15 in 2006	230 (5,027)	15	0	15	15
Individuals age 16 in 2006	176 (5,027)	16	0	16	16
Individuals age 17 in 2006	129 (5,027)	17	0	17	17
Individuals age 18-29 in 2006	2,544 (5,027)	23	3.402	18	29
District municipalities	7,718			1	34
Wealth (sample of interest$_1$)	5,027	1.916	0.860	1	3
Low income	2,090 (5,027)				
Middle income	1,270 (5,027)				
High income	1,667 (5,027)				
Ethnicity (sample of interest$_1$)	5,027	1.198	0.601	1	5
Black/African	4,502 (5,027)				
White	92 (5,027)				
Coloured	396 (5,027)				
Indian/Asian	36 (5,027)				
Literacy (sample of interest$_1$)	5,024	2.934	0.275	1	3
Cannot read	36 (5,024)				
Can read a little	257 (5,024)				
Can read	4,731 (5,024)				

Native language (sample of interest[1])	5,027	2.744	0.619	1	3
English	485 (5,027)				
Afrikaans	316 (5,027)				
Other	4,226 (5,027)				
Panel B: LSE					
LSE intensity municipalities	7,718	0.485	0.500	0	1
Log(LSE) index	8,737	0.933	0.348	2.38e-07	1.493
Panel C: SRB					
Log(SRB) index	8,737	0.474	0.450	0	1.449
Panel D: HIV					
HIV status	8,700	0.794	0.404	0	1

[1]*Sample of interest is individuals with no, little or full exposure to the LSE programs. The values within the parentheses corresponds to the total number of observations in the sample of interest.*

4 Methodological approach and design

This section presents the research design and statistical technique utilised. It also introduces the equations for the regression models as well as a description of important calculations to bear in mind when reading the tables presenting the output of the regressions.

4.1 Statistical technique

Considering that both the 2005 LSE program and the 2011 LSE program were designed to be full coverage programs, randomisation is not possible. However, a quasi-experimental evaluation design, such as the Difference-in-difference (DiD) approach, is suitable (Columbia University Mailman School of Public Health, 2019). The method requires data from before and after the intervention (Deaton, A. 2013:165), and since the cohort of exposure and the control group cannot be observed during the same period, DiD takes a parallel trends assumption. This implies that if there is no treatment, then the differences between the cohorts are constant over time. This is the central assumption of DiD, and if it is breached, the estimation of the effect will be biased (Columbia University Mailman School of Public Health, 2019). To assess whether it is likely that the parallel trends assumption holds, we present six control regressions. In all regressions a cohort aged 18 to 21 in 2006 (they were 28 to 31 in 2016) are used as "the cohort of exposure", whilst a cohort aged 22 to 25 in 2006 (they were 32 to 35 in 2016) are used as "the control group"[21].

Additionally, DiD limits the risk of getting biased results because of general issues such as homoscedasticity, autocorrelation, multicollinearity, the presence of influential observations, endogeneity and spurious relationships. Nonetheless, a few statistical tests are conducted to further mitigate the impact of these problems. First, this research tests for the normality of residuals in all models, since normal distributed errors generate more accurate t- and F-values (Mehmetoglu and Jakobsen, 2017:151-153). It was noted that all models have some problem with non-normal distributed residuals. However, this is not of big concern, since non-normal distributed errors are mainly problematic in models with small samples (Mehmetoglu and Jakobsen, 2017:151-153). Nonetheless, two of the three main dependent variables, that is the variables measuring the level of LSK and the level of SRB, are logged in an attempt to conform the models into normal distributions (Mehmetoglu & Jakobsen, 2017:326-331). A slightly more normalised distribution is ensured when using the logged transformation of

[21] See table 6 in the result section for the results of the control regressions.

LSK and SRB. Thus, the logged LSK and SRB are used throughout this research. The results of all nine tests with the logged transformed LSK and SRB variables are presented in figure A1 in appendix.

Secondly, all models are expected to have some issues with both heteroscedasticity and autocorrelation. The former is based on the assumption that the error term has constant variance, whilst the latter requires that the standard errors are not correlated. Breaching any of these two assumptions might result in biased standard errors. The diagnostic tests carried out to check for heteroscedasticity and autocorrelation yield the anticipated results. Consequently, it is concluded that there is a situation of heteroscedasticity and autocorrelation (Mehmetoglu & Jakobsen, 2017:149-151). In order to account for these issues, robust and clustered standard errors are used. By using the robust option, the assumption of constant variance of the error term is relaxed, whilst the clustered option mitigates the problem of correlated standard errors. The clustering is made on the basis of municipalities; thus, the standard errors are calculated based on the number of municipalities, instead of the total number of observations (Mehmetoglu & Jakobsen, 2017:234-235).

The assumption of no-multicollinearity implies that the independent variables must be uncorrelated. Otherwise, there is a risk of estimating too low standard errors. However, considering that all models include one or more interaction terms, the assumption of absence of multicollinearity is likely to be violated (Mehmetoglu & Jakobsen, 2017:146-147). Nevertheless, after checking the variance inflation factor (VIF), it was concluded that multicollinearity does not pose a big problem in any of the models. The estimated mean VIF-scores for each model is presented in table A6 in appendix. Additionally, this research also controls for the presence of influential observations, i.e. outliers that would substantially influence the outcome if they were excluded from the analysis. By using the Cook's distance estimate it is possible to check the influence of each observation on the model as a whole. It is noted that a few observations in each of the estimated models exert a bit more influence than others (Mehmetoglu & Jakobsen, 2017:153-157). However, Mehmetoglu and Jakobsen (2017:153-157) suggest a cut-off value of 1 and none of the observations in any of the models exceeds a Cook's distance value of 1. Consequently, it is concluded that influential observations should not constitute a problem. Figure A2 in appendix presents the results from the Cook's distance measurement.

Furthermore, the issues of endogeneity and spurious relationships are connected, where the former entails that one or more of the independent variables correlates with the error term (Aneshensel, 2013:85-95), whilst the latter implies that the observed relationship between the

independent variables and the dependent variable is impacted by an unobserved factor (Aneshensel, 2013:199-222). If either of the assumptions are breached, there is a risk of generating omitted variable bias. In order to mitigate the risk of endogeneity and spuriousness, this study includes several control variables to rule out counterarguments (Aneshensel, 2013:85-95). In order to alleviate the risk of omitted variable bias further, this study also utilises fixed effects. Thus, it is possible to control for factors that are constant across individuals, whilst also enabling the model to capture the variation between individuals within municipalities (Mehmetoglu & Jakobsen, 2017:240-249).

Finally, to account for disproportionate sampling, this research includes analytical weights in all regressions. In the models where the level of LSK and the level of SRB are used as dependent variables, the women's individual sample weight is utilised, whereas in the models where HIV status is the dependent variable, the HIV sample weight is used instead. The weighting is done using the probability of sampling weights (pweight) command since both the individual sample weight and the HIV sample weight were set to the inverse of the probability of an observation to be included in the sample. Thus, by including analytical weights, it is possible to generalise the results to the whole population (Mehmetoglu & Jakobsen, 2017:331-333). In the following section the execution of the statistical method is described step-by-step.

4.2 Execution of statistical method

This study follows the model used by Duflo (2001). In her research she divides regions in "high program" and "low program" regions based on the intensity of an education program. She then compares the educational attainment and the wages of individuals who had little or no exposure to the program to those individuals who were fully exposed to the program in each region. Similar to Duflo (2001), this study employs a DiD approach by first comparing the level of life skills knowledge, sexual risk behaviour and HIV prevalence of individuals with *full* exposure to the LSE programs (they were 6 to 9 in 2006) to those with no exposure to the programs (they were 18 to 21 in 2006). Secondly, we compare the outcomes of individuals with *little* or *full* exposure (they were 6 to 17 in 2006) to those with no exposure (they were 18 to 29 in 2006).

A few preparations were made before the actual analysis could take place. Initially, a selection of variables that capture at least one of the eleven aspired outcomes of the 2005

LSE[22] is made from first the 1998 dataset and then from the 2016 dataset[23]. Thereafter, the selected variables are recoded and made sure to move in tandem, i.e. in the same direction. A similar selection and recoding is done for relevant variables measuring sexual risk behaviour[24]. In the following step, three different indices are created (*1998 initial LSK index, 2016 LSK index* and *2016 SRB index*) by using multiple correspondence analysis (MCA). MCA is a favourable method to use when creating indices from categorical variables since it allows for the analysis of patterns between several nominal variables (Abdi & Valentin, 2007:1-2). The *1998 initial LSK*, is then used to separate municipalities into "high initial LSK" and "low initial LSK" municipalities. The purpose of this division is to enable the first difference, that is the estimation of initial level of life skills knowledge, from 1998 to 2016. This difference is possible since both LSE programs were designed to have full national coverage and because it is reasonable to assume that LSK differs across municipalities.

In the last preparatory step, individuals are separated into different cohorts based on their level of exposure to the LSE programs. The groups of interest consist of individuals with *little* or *full* exposure to the LSE programs[25], whilst the control groups are comprised of individuals with no exposure to the programs[26]. Hence, the comparison of the change in outcome of the main group of interest with the change in outcome of the baseline group constitutes the second difference.

The main analysis is divided into three steps. First, we estimate Equation 1: Unconditional model. This regression is the simplest model estimating the effect of the LSE programs on three different outcomes (Y_{ijk}); (1) the logged level of life skills knowledge of individual *i* studying in municipality *j* in year *k*, (2) the logged level of sexual risk behaviour of individual *i* studying in municipality *j* in year *k*, and (3) HIV prevalence of individual *i* studying in municipality *j* in year *k*. Moreover, *c* is a constant, α_j is initial LSK municipality fixed effect, P_j is a dummy implying the intensity of initial life skills knowledge in municipality of schooling, T_i is a dummy denoting whether the individual belongs to the young cohort of exposure and β_k is a cohort of birth fixed effect.

[22] See the background section for the aspired outcomes of the 2005 LSE program.
[23] The dataset used in the study is described in section *3.1 Dataset.*
[24] See the variable section for an overview of the variables used in this research and to get an understanding of how these have been adjusted to fit the framework of analysis.
[25] All individuals included in the cohort of exposure have at least 1 year of exposure to the LSE programs. Thus, individuals aged 6 to 17 with no exposure are excluded from the group.
[26] First, the effect of the programs is compared between individuals with *full* exposure (they were 6 to 9 in 2006) to those with no exposure (they were 18 to 21 in 2006). Second, the effect of the programs is compared between individuals with *little* or *full* exposure (they were 6 to 17 in 2006) to those with no exposure (they were 18 to 29 in 2006).

Equation 1: Unconditional model

$$Y_{ijk} = c + \alpha_j + \beta_k + (P_j T_i) \cdot \gamma + \varepsilon_{ijk}$$

Second, we estimate Equation 2: Short model. The short model is similar to the unconditional model, however, it includes variables controlling for an individual's level of wealth (W_i), native language (LA_i), level of literacy (LI_i) as well as ethnicity (E_i).

Equation 2: Short model

$$Y_{ijk} = c + \alpha_j + \beta_k + (P_j T_i) \cdot \gamma + \delta \cdot W_i + \lambda \cdot LA_i + \mu \cdot LI_i + \omega \cdot E_i + \varepsilon_{ijk}$$

Finally, Equation 3: Long model, is estimated. The main difference between the long and the short model is that the effect of the programs is estimated on each individual age group between 6 to 17-year old by using 12 year-of-birth dummies. Thus, making it possible to distinguish the impact of the programs on individuals with different levels of exposure.

Equation 3: Long model

$$Y_{ijk} = c + \alpha_j + \sum_{l=6}^{17} (d_{il}) + \sum_{l=6}^{17} (P_j d_{il}) \cdot \gamma_l + \delta \cdot W_i + \lambda \cdot LA_i + \varrho \cdot LI_i + \omega \cdot E_i + \varepsilon_{ijk}$$

We expect that individuals with exposure to the LSE programs will have higher LSK compared to those with no exposure to the programs. It is also hypothesised that the effect on the level of LSK will decrease with increasing ages, since older individuals in the cohort of exposure have had less exposure to the programs compared to the younger ones. Moreover, it is anticipated that individuals who have high LSK, will have low level of SRB and similarly, low HIV prevalence rates. Thus, we expect that the effect on the level of LSK will gradually decrease with older ages, whilst the effect on the level of SRB and HIV prevalence will increase. Finally, it is also anticipated that individuals from "low initial LSK" municipalities will on average have a bigger increase in LSK as well as a bigger decrease in SRB and HIV prevalence, compared to individuals from "high initial LSK" municipalities. This assumption is based on the belief that individuals with an initial low LSK will have gained more

knowledge from the LSE programs, considering that they did not have a lot of knowledge from before. In order to distinguish differences between the two cohorts across as well as within "high initial LSK" and "low initial LSK" municipalites, we use the following calculations.

Calculation 1: Difference between young cohort of exposure and cohort of control in a "high initial LSK" municipality

$$D_1 = \left(c + \alpha_j + \beta_k + \left(P_j T_i\right) \cdot \gamma\right) - \left(c + \alpha_j\right)$$

Calculation 2: Difference between young cohort of exposure and cohort of control in a "low initial LSK" municipality

$$D_2 = \left(c + \beta_k\right) - \left(c\right)$$

Calculation 3: Difference between young cohort of exposure in a "high initial LSK" and cohort of exposure in a "low initial LSK" municipality

$$D_3 = \left(c + \alpha_j + \beta_k + \left(P_j T_i\right) \cdot \gamma\right) - \left(c + \beta_k\right)$$

Calculation 4: Difference between young cohort of control in a "high initial LSK" and cohort of control in a "low initial LSK" municipality

$$D_4 = \left(c + \alpha_j\right) - \left(c\right)$$

Calculation 5: Difference in difference

$$DiD = (D_1 - D_2) - (D_3 - D_4)$$

We obtain the DiD by considering the interaction term, however, this can also be obtained by calculating the difference between the two cohorts and the difference between "high initial LSK" and "low initial LSK" municipalities.

5 Results

The purpose of this section is to present the results obtained from the estimations. First, we present the control regressions to test whether the parallel trend assumption holds. Second, the results from the unconditional and short models are provided. Lastly, we present the results of the long models. These findings are analysed, discussed and related back to theory in the subsequent section, which is followed by a short conclusion.

5.1 Testing the impact of a placebo programs in pre-programs cohorts

Table 6 presents the results of the six regressions controlling for the parallel trend assumption estimated using Equation 1 and 2. The main difference between the control regressions and the regression of interest is that the young cohort of exposure (individuals aged 6 to 9 in 2006) is replaced by a cohort aged 18 to 21 in 2006, and the cohort of control is replaced by a cohort aged 22 to 25 in 2006. Thus, neither of the included cohorts in the control regressions have had exposure to the LSE programs[27].

Table 6: Control regressions: Unconditional and short model

		lnLSK			lnSRB			HIV		
		High (1)	Low (2)	Diff (3)	High (4)	Low (5)	Diff (6)	High (7)	Low (8)	Diff (9)
Panel A: Unconditional model										
Age 18 to 21 in 2006	**(1)**	1.02	1.01	0.01	0.30	0.30	0.00	0.78	0.77	0.01
		(0.15)	(0.06)	(0.09)	(0.14)	(0.06)	(0.08)	(0.19)	(0.08)	(0.11)
Age 22 to 25 in 2006	**(2)**	1.10	1.03	0.07	0.33	0.36	-0.03	0.72	0.79	-0.07
		(0.07)	(0.03)	(0.04)	(0.07)	(0.03)	(0.04)	(0.10)	(0.04)	(0.06)
Difference	**(3)**	-0.08	-0.02	-0.06	-0.03	-0.06	0.03	0.06	-0.02	0.08
		(0.08)	(0.03)	(0.05)	(0.07)	(0.03)	(0.05)	(0.09)	(0.04)	(0.05)
Panel B: Short model										
Age 18 to 21 in 2006	**(4)**	1.46	1.48	-0.02	0.23	0.25	-0.02	0.98	1.00	-0.02
		(0.24)	(0.14)	(0.10)	(0.20)	(0.11)	(0.09)	(0.25)	(0.15)	(0.10)
Age 22 to 25 in 2006	**(5)**	1.54	1.49	0.05	0.27	0.31	-0.04	0.92	1.01	-0.09
		(0.16)	(0.11)	(0.05)	(0.13)	(0.08)	(0.05)	(0.17)	(0.11)	(0.06)
Difference	**(6)**	-0.08	-0.01	-0.07	-0.04	-0.06	0.03	0.06	-0.01	0.08
		(0.08)	(0.03)	(0.05)	(0.07)	(0.03)	(0.05)	(0.08)	(0.04)	(0.04)

*Comment: Short model includes control variables for literacy, wealth, language and ethnicity. The SRB variable is coded so that a high value of SRB indicates low sexual risk behavior. HIV variable is coded so that a value of 1 corresponds to the individual being HIV negative, whilst a value of 0 denotes that the individual is HIV positive. Clustering is made on the basis of municipalities. Standard errors in parenthesis. *p<0.05, **p<0.01, ***p<0.001*

[27] See section *4.1 Statistical Technique* for more information on the control regressions.

Row 3 and 6 in columns 1, 2, 4, 5, 7 and 8 presents the differences (D_1 and D_2) obtained from Calculation 1 and 2. The estimates in column 3, 6 and 9 row 1, 2, 4 and 5 demonstrates the difference (D_3 and D_4) obtained from Calculation 3 and 4. Lastly row 3 and 6 in column 3, 6 and 9 estimates the difference in difference (DiD), i.e. the effect of the programs. In order to not violate the parallel trend assumption, the estimated DiD of the control regressions should be close to 0. As can be noted in panel A and B, column 3, 6 and 9, the estimated effect (DiD) in all six control regressions are almost 0. In addition, the impact of the programs is statistically insignificant in all estimations. Thus, it is possible to conclude that the parallel trend assumption holds, i.e. the models are not driven by any inappropriate identification assumptions.

5.2 Effect of the LSE programs on level of life skills knowledge, level of sexual risk behaviour and HIV prevalence by cohort and initial level of LSK

Table 7 demonstrates the results of the regressions of interest using equation 1 and 2. The estimates in table 7 should be interpreted in the same way as in table 6.

Table 7: Regressions of interest - Equation 1: Unconditional model and Equation 2: Short model

		lnLSK			lnSRB			HIV		
		High (1)	Low (2)	Diff (3)	High (4)	Low (5)	Diff (6)	High (7)	Low (8)	Diff (9)
Panel A: Unconditional model										
Age 6 to 9 in 2006	**(1)**	0.73	0.68	0.05	0.89	0.90	-0.01	0.88	0.86	0.02
		(0.16)	(0.06)	(0.10)	(0.15)	(0.05)	(0.10)	(0.18)	(0.07)	(0.11)
Age 18 to 21 in 2006	**(2)**	1.02	1.01	0.01	0.30	0.30	0.00	0.78	0.77	0.01
		(0.08)	(0.03)	(0.05)	(0.03)	(0.01)	(0.02)	(0.08)	(0.03)	(0.05)
Difference	**(3)**	-0.29	-0.33	0.04	0.59	0.60	-0.01	0.10	0.09	0.00
		(0.08)	(0.03)	(0.05)	(0.12)	(0.04)	(0.08)	(0.10)	(0.04)	(0.06)
Panel B: Short model										
Age 6 to 9 in 2006	**(4)**	1.30	1.27	0.03	0.79	0.86	-0.07	0.99	0.99	0.00
		(0.29)	(0.19)	(0.10)	(0.32)	(0.22)	(0.10)	(0.24)	(0.14)	(0.10)
Age 18 to 21 in 2006	**(5)**	1.58	1.60	-0.02	0.21	0.24	-0.03	0.90	0.90	0.00
		(0.21)	(0.16)	(0.05)	(0.20)	(0.19)	(0.01)	(0.15)	(0.10)	(0.05)
Difference	**(6)**	-0.28	-0.33	0.05	0.58	0.62	-0.03	0.09	0.09	0.00
		(0.08)	(0.03)	(0.05)	(0.12)	(0.03)	(0.08)	(0.09)	(0.04)	(0.05)

*Comment: Short model includes control variables for literacy, wealth, language and ethnicity. The SRB variable is coded so that a high value of SRB indicates low sexual risk behavior. HIV variable is coded so that a value of 1 corresponds to the individual being HIV negative, whilst a value of 0 denotes that the individual is HIV positive. Clustering is made on the basis of municipalities. Standard errors in parenthesis. *p<0.05, **p<0.01, ***p<0.001*

The table suggest that the average effect of the programs is 4% to 5% on the level of life skills knowledge, negative 1% to 3% on the level of sexual risk behaviour and 0 percentage points on probability of having a positive HIV status. In both models, and in both types of initial LSK municipalities, the average LSK, SRB and HIV prevalence decreased over time. Additionally, the decrease in level of LSK and SRB is bigger in "low initial LSK" municipalities, whilst the decrease in HIV prevalence is as big in both types of municipalities. The third difference (D_3) in column 3 row 1 and 4 suggests that the average level of life skills knowledge is 3% to 5% higher among young individuals who attended school in a "high initial LSK" municipality, compared to young individuals who attended school in a "low initial LSK" municipality. Similarly, according to column 9 row 1 and 4, young individuals in "high initial LSK" municipalities have on average 2 percentage points higher HIV prevalence than young individuals in "low initial LSK" municipalities. Furthermore, D_3 in column 6 row 1 and 4 indicate that the young individuals in "low initial LSK" municipalities have on average 1% to 7% lower level of sexual risk behaviour than young individuals in "high initial LSK" municipalities. Finally, by comparing the estimates of the unconditional and the short model, it can be the noted that the inclusion of controls increased the effect size slightly. Regardless, the estimates are statistically insignificant in all models, thus, it can be concluded that the LSE programs have not had the expected effect on the level of life skills knowledge, level of sexual risk behaviour or HIV prevalence.

5.3 Effect of the LSE programs on level of life skills knowledge, level of sexual risk behaviour and HIV prevalence by year of birth and initial level of LSK

The results of Equation 3: Long model, are presented and interpreted below. Initially, the result of the regression estimating the effect of the LSE programs on level of life skills knowledge is presented. This if followed by a demonstration of the result obtained from the estimation of the effect on sexual risk behaviour. Finally, the outcome of the regression estimating the effect on HIV prevalence is presented.

Table 8: Effect of the programs on level of LSK knowledge - coefficients of the interaction between year-of-birth dummies indicating age in 2006 and high initial LSK intensity municipality

	Dependent variable: Log (LSK)				
	(1)	(2)	(3)	(4)	(5)
Age in 2006					
6	-1.043***	-0.379***	-0.376***	-0.376***	-0.376***
	(-13.21)	(-13.15)	(-13.09)	(-12.86)	(-12.99)
7	-1.033***	-0.371***	-0.365***	-0.361***	-0.363***
	(-20.37)	(-23.82)	(-22.78)	(-22.85)	(-23.33)
8	-1.006***	-0.362***	-0.359***	-0.359***	-0.359***
	(-19.79)	(-16.27)	(-16.16)	(-15.73)	(-15.98)
9	-0.856***	-0.304***	-0.305***	-0.303***	-0.305***
	(-12.21)	(-12.27)	(-13.09)	(-12.51)	(-12.84)
10	-0.745***	-0.258***	-0.257***	-0.255***	-0.258***
	(-12.69)	(-11.39)	(-11.15)	(-11.53)	(-11.87)
11	-0.766***	-0.270***	-0.267***	-0.269***	-0.268***
	(-10.48)	(-10.09)	(-9.84)	(-10.54)	(-10.38)
12	-0.783***	-0.270***	-0.271***	-0.267***	-0.265***
	(-6.94)	(-6.66)	(-6.58)	(-6.62)	(-6.44)
13	-0.350*	-0.125*	-0.123	-0.123*	-0.121
	(-2.05)	(-2.06)	(-2.01)	(-2.09)	(-2.03)
14	-0.367**	-0.131**	-0.129**	-0.129**	-0.126**
	(-3.17)	(-3.37)	(-3.39)	(-3.24)	(-3.21)
15	-0.516***	-0.178***	-0.179***	-0.176***	-0.175***
	(-4.21)	(-4.17)	(-4.13)	(-4.15)	(-4.06)
16	-0.586***	-0.207***	-0.206***	-0.205***	-0.209***
	(-11.10)	(-12.44)	(-11.84)	(-11.87)	(-12.19)
17	-0.239*	-0.0772	-0.0799*	-0.0895*	-0.0842*
	(-2.12)	(-2.03)	(-2.14)	(-2.42)	(-2.20)
Initial LSK intensity	0.0102	0.0344	0.0296	0.0219	0.0200
	(0.94)	(0.93)	(0.80)	(0.59)	(0.53)
Age in 2006 × LSK intensity					
6	-0.0768	-0.0179	-0.0226	-0.0318	-0.0251
	(-0.69)	(-0.44)	(-0.57)	(-0.69)	(-0.54)
7	-0.0641	-0.0137	-0.0185	-0.0180	-0.0127
	(-0.66)	(-0.42)	(-0.56)	(-0.54)	(-0.36)
8	0.222	0.0891	0.0831	0.0762	0.0863
	(1.23)	(1.38)	(1.24)	(1.19)	(1.37)
9	-0.114	-0.0287	-0.0274	-0.0309	-0.0251
	(-1.04)	(-0.78)	(-0.76)	(-0.80)	(-0.64)
10	-0.115	-0.0317	-0.0360	-0.0345	-0.0293
	(-1.22)	(-0.98)	(-1.12)	(-1.15)	(-0.96)
11	0.0124	0.0200	0.0165	0.0252	0.0266
	(0.11)	(0.54)	(0.43)	(0.71)	(0.73)
12	0.112	0.0455	0.0522	0.0558	0.0551
	(0.80)	(0.94)	(1.10)	(1.20)	(1.15)
13	-0.279	-0.0838	-0.0879	-0.0887	-0.0789
	(-1.42)	(-1.23)	(-1.27)	(-1.29)	(-1.11)
14	-0.300	-0.0971	-0.0978	-0.0921	-0.0901
	(-1.78)	(-1.66)	(-1.66)	(-1.52)	(-1.48)
15	0.0243	0.0153	0.0143	0.0183	0.0209
	(0.15)	(0.28)	(0.26)	(0.34)	(0.39)
16	0.406**	0.147**	0.146**	0.145**	0.153**
	(2.77)	(3.09)	(3.03)	(3.15)	(3.36)
17	0.0202	0.00274	0.00567	0.0111	-0.00508
	(0.12)	(0.05)	(0.09)	(0.17)	(-0.07)
Control variables					
Literacy (base line: Cannot read)					
Can read a little		-0.182	-0.176	-0.159	-0.152
		(-1.93)	(-1.89)	(-1.86)	(-1.77)

Can read		-0.146 (-1.72)	-0.145 (-1.72)	-0.132 (-1.70)	-0.123 (-1.56)
Wealth (base line: low income)					
Middle income			0.0114 (0.67)	0.00711 (0.41)	0.00833 (0.48)
High income			0.0333 (1.58)	0.0153 (0.63)	0.00571 (0.24)
Language (base line: Other languages)					
English				0.0901* (2.59)	0.0415 (0.95)
Afrikaans				0.0895** (3.00)	-0.00995 (-0.14)
Ethnicity (base line: White)					
Black/African					-0.195* (-2.61)
Coloured					-0.109*** (-3.76)
Indian/Asian					-0.0630 (-0.75)
Constant	2.978*** (39.34)	1.177*** (12.94)	1.164*** (12.98)	1.149*** (13.74)	1.337*** (10.72)
F-statistic	3.11	3.28	3.44	3.69	3.90
R^2	0.15	0.16	0.16	0.17	0.17
N	4,648	4,645	4,645	4,645	4,645
t statistics in parentheses *** p<0.05, ** p<0.01, *** p<0.001**					

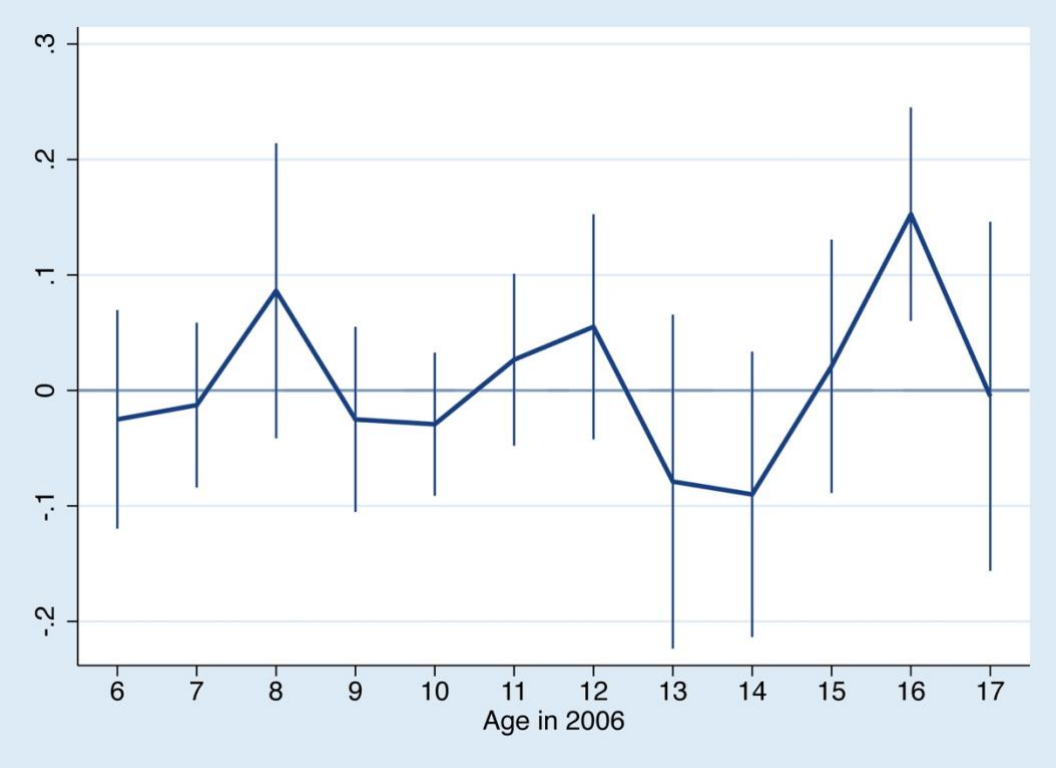
Comment: Column (1) estimates the effect of life skills knowledge without controls. Column (2) adds a variable controlling for literacy, column (3) adds control for wealth, column (4) adds a variable controlling for home language and column (5) adds control for ethnicity. Clustering is made on the basis of municipalities.
F-statistics test the hypotheses that the coefficients of interaction between year of birth dummies and the LSK intensity are jointly 0,

Table 8 provide for the results from the estimations of the effect of the programs on the level of life skills knowledge. The coefficients are statistically insignificant for all interaction terms except from the one measuring the effect on 16-years-old. Seemingly, the effect on 16 years old ranges between 15% in the model including all controls but ethnicity (column 4), to 41% in the model without controls (column 1). In the model including all control variables (column 5), the estimated effect is 15% on the level of life skills knowledge. Moreover, no clear pattern for whether the effect increases or decreases with age can be distinguished. Although, not statistically significant, figure 5 clearly illustrates the inconsistent effect of the programs on each individual age group.

Nevertheless, the majority of the variables capturing the average level of life skills knowledge are statistically significant. Seemingly, the control group, i.e. older individuals, have in general higher LSK compared to the cohort of exposure, and within the cohort of exposure, younger individuals tend to have lower average of LSK compared to older individuals. Furthermore, the only statistically significant controls are noted in column 4 and 5. In column 4, it is estimated that having English or Afrikaans as home language is

associated with a 9% higher LSK compared to having a local language as home language. However, when controlling for ethnicity (column 5), English and Afrikaans loses their statistically significance. Instead, the only statistically significant factors among the controls are black/African and Coloured. It is estimated that black/Africans and Coloured have 20% and 11% lower LSK, respectively, compared to white individuals.

Figure 5: Effect of the programs on level of LSK



Comment: The coefficients of the interaction between year-of-birth dummies indicating age in 2006 and high LSK intensity municipality for the LSK model.

Table 9: Long SRB model - coefficients of the interaction between year-of-birth dummies indicating age in 2006 and high initial LSK intensity municipality

	Dependent variable: Log(SRB)				
	(1)	(2)	(3)	(4)	(5)
Age in 2006					
6	0.889***	0.890***	0.895***	0.893***	0.894***
	(13.52)	(13.49)	(13.42)	(13.47)	(13.42)
7	0.645***	0.646***	0.657***	0.661***	0.660***
	(7.08)	(7.10)	(7.22)	(7.24)	(7.24)
8	0.493***	0.493***	0.499***	0.500***	0.502***
	(6.14)	(6.14)	(6.27)	(6.23)	(6.29)
9	0.389***	0.390***	0.389***	0.390***	0.389***
	(10.07)	(10.05)	(10.89)	(10.37)	(10.18)
10	0.197**	0.197**	0.199**	0.201**	0.198**
	(3.39)	(3.39)	(3.44)	(3.48)	(3.50)
11	0.0717	0.0717	0.0770	0.0756	0.0770
	(0.97)	(0.97)	(1.08)	(1.05)	(1.07)
12	0.133*	0.134*	0.132*	0.136*	0.138*
	(2.42)	(2.40)	(2.41)	(2.48)	(2.52)
13	0.0233	0.0214	0.0256	0.0263	0.0280
	(0.44)	(0.40)	(0.49)	(0.49)	(0.52)
14	-0.00813	-0.00774	-0.00426	-0.00471	-0.00296
	(-0.22)	-0.00774	(-0.12)	(-0.13)	(-0.08)
15	0.0166	0.0165	0.0153	0.0176	0.0188
	(0.44)	(0.45)	(0.40)	(0.47)	(0.51)
16	-0.115	-0.114	-0.112	-0.113	-0.115
	(-1.43)	(-1.41)	(-1.39)	(-1.35)	(-1.34)
17	-0.0629**	-0.0627**	-0.0680**	-0.0813***	-0.0772**
	(-3.31)	(-3.28)	(-3.55)	(-3.74)	(-3.55)
Initial LSK intensity	0.000811	0.000460	-0.00914	-0.0157	-0.0159
	(0.05)	(0.03)	(-0.50)	(-0.78)	(-0.77)
Age in 2006 × LSK intensity					
6	-0.105	-0.105	-0.114	-0.122	-0.117
	(-0.64)	(-0.64)	(-0.70)	(-0.71)	(-0.70)
7	0.106	0.105	0.0955	0.0978	0.102
	(0.86)	(0.85)	(0.76)	(0.76)	(0.80)
8	0.0124	0.0123	0.000606	-0.00532	-0.000694
	(0.09)	(0.09)	(0.00)	(-0.04)	(-0.01)
9	-0.113	-0.114	-0.111	-0.114	-0.107
	(-1.39)	(-1.40)	(-1.42)	(-1.48)	(-1.37)
10	0.0634	0.0638	0.0555	0.0567	0.0599
	(0.78)	(0.78)	(0.70)	(0.71)	(0.75)
11	-0.0702	-0.0690	-0.0760	-0.0670	-0.0671
	(-0.89)	(-0.87)	(-0.98)	(-0.86)	(-0.86)
12	-0.188**	-0.188**	-0.175*	-0.171*	-0.172*
	(-2.88)	(-2.87)	(-2.65)	(-2.57)	(-2.60)
13	0.0455	0.0476	0.0396	0.0329	0.0352
	(0.60)	(0.62)	(0.51)	(0.41)	(0.45)
14	0.0520	0.0522	0.0509	0.0588	0.0588
	(0.69)	(0.69)	(0.66)	(0.79)	(0.78)
15	0.0347	0.0361	0.0343	0.0397	0.0403
	(0.48)	(0.51)	(0.48)	(0.56)	(0.57)
16	0.0436	0.0434	0.0425	0.0446	0.0518
	(0.51)	(0.51)	(0.51)	(0.51)	(0.59)
17	0.184*	0.185*	0.191*	0.202**	0.181**
	(2.36)	(2.37)	(2.57)	(2.88)	(2.76)
Control variables					
Literacy (base line: Cannot read)					
Can read a little		0.0164	0.0276	0.0376	0.0340
		(0.16)	(0.28)	(0.38)	(0.34)

Can read		-0.00688	-0.00514	0.00220	-0.0000127
		(-0.08)	(-0.06)	(0.02)	(-0.00)
Wealth (base line: low income)					
Middle income			0.0220	0.0176	0.0188
			(0.77)	(0.62)	(0.65)
High income			0.0656**	0.0466	0.0400
			(2.77)	(1.99)	(1.87)
Language (base line: Other languages)					
English				0.0678**	0.0359
				(2.76)	(0.85)
Afrikaans				0.130***	0.0981
				(3.94)	(1.50)
Ethnicity (base line: White)					
Black/African					-0.0873
					(-1.06)
Coloured					-0.0647
					(-0.95)
Indian/Asian					0.106
					(0.90)
Constant	0.326***	0.331***	0.306**	0.296**	0.387**
	(26.12)	(3.61)	(3.47)	(3.34)	(2.97)
F-statistic	4.51	4.41	4.76	5.10	4.90
R^2	0.24	0.24	0.24	0.25	0.25
N	4,648	4,645	4,645	4,645	4,645
t statistics in parentheses					
*** p<0.05, ** p<0.01, *** p<0.001**					

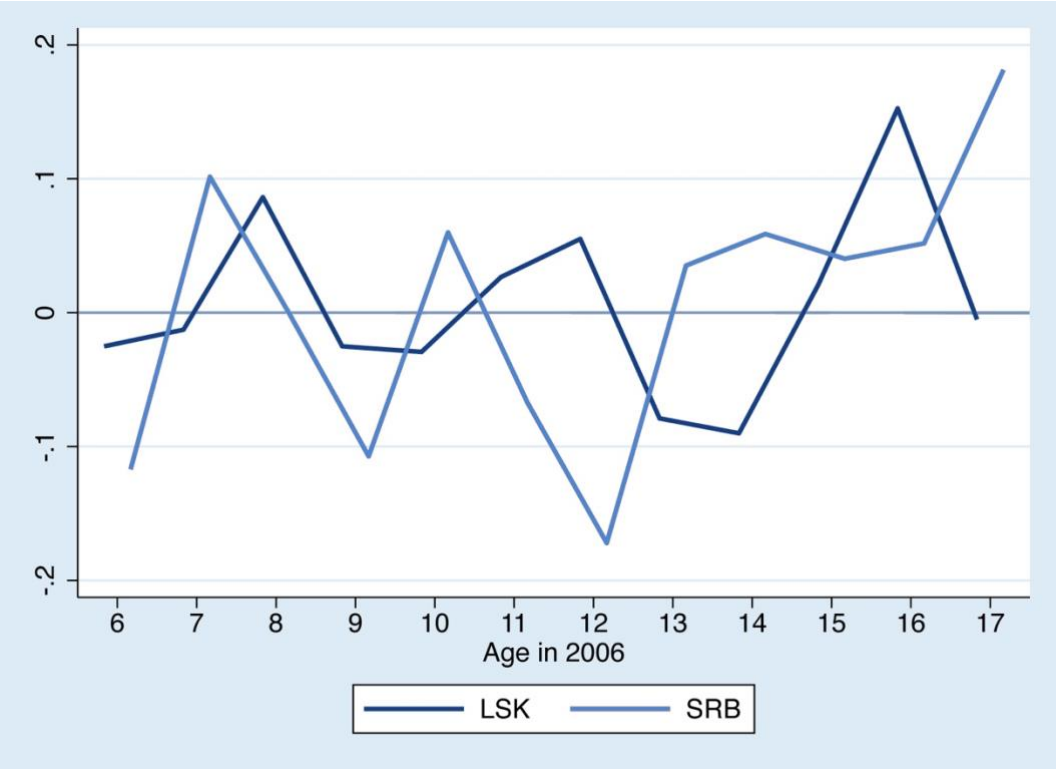
Comment: Column (1) estimates the effect of life skills knowledge without controls. Column (2) adds a variable controlling for literacy, column (3) adds control for wealth, column (4) adds a variable controlling for home language and column (5) adds control for ethnicity. The SRB variable is coded so that a high value of SRB indicates low sexual risk behavior. Clustering is made on the basis of municipalities. F-statistics test the hypotheses that the coefficients of interaction between year of birth dummies and the LSK intensity are jointly 0,

The results from the regressions estimating the effect of the programs on sexual risk behaviour is presented in table 9. The majority of the estimates measuring the average level of SRB are statistically significant. It can be noted that younger individuals tend to have on average lower sexual risk behaviour compared to older individuals, both within the cohort of exposure and when comparing to the control group. In contrast, the majority of the coefficients estimating the effect of the programs are statistically insignificant. However, the effect is statistically significant on individuals aged 12 and 17. Among 12-years-old, the effect ranges between negative 17% (column 4 and 5) to negative 19% (column 1 and 2), whilst for 17-years-old, the effect ranges between 18% (column 1 and 5) to 20% (column 4). Figure 6 shows the effect of the programs on level of LSK as well as on level of SRB among the 12 different age groups. As can be noted, no clear patterns can be distinguished.

Similarly, to the model estimating the effect on LSK, the gradually inclusion of control variables does not have any clear impact on the effect size. Moreover, only the variables controlling for high income takers in column 3 and English and Afrikaans as home languages in column 4 are statistically significant. The results indicate that high income takers have 6%

lower SRB compared to low income takers. Furthermore, individuals who have either English or Afrikaans as their home language have 7% and 13% lower levels of SRB, respectively, compared to individuals who have local languages as their home language. However, all controls lose their statistically significance in the full model (column 5).

Figure 6: Effect of the programs on level of LSK and level of SRB



Comment: The coefficients of the interaction between year-of-birth dummies indicating age in 2006 and high LSK intensity municipality for each model, LSK and SRB.

Table 10: Long HIV model

	Dependent variable: HIV				
	(1)	(2)	(3)	(4)	(5)
Age in 2006					
6	0.0917	0.0906	0.0992	0.0971	0.0944
	(1.80)	(1.78)	(1.88)	(1.86)	(1.80)
7	0.0623	0.0631	0.0812	0.0847	0.0826
	(1.17)	(1.20)	(1.57)	(1.61)	(1.55)
8	0.122**	0.121*	0.130**	0.131**	0.129**
	(2.77)	(2.73)	(2.94)	(2.98)	(2.94)
9	0.0417	0.0406	0.0351	0.0355	0.0333
	(0.87)	(0.85)	(0.67)	(0.71)	(0.69)
10	0.128***	0.128***	0.126***	0.128***	0.123***
	(3.85)	(3.86)	(3.83)	(3.97)	(3.82)
11	0.0551	0.0552	0.0647	0.0602	0.0642
	(0.70)	(0.70)	(0.88)	(0.75)	(0.85)
12	0.0679	0.0667	0.0609	0.0647	0.0680
	(1.49)	(1.47)	(1.33)	(1.42)	(1.48)
13	0.0889	0.0914*	0.0967*	0.0965*	0.100*
	(1.91)	(2.05)	(2.15)	(2.18)	(2.20)
14	0.104**	0.104**	0.107**	0.108**	0.110**
	(3.01)	(2.99)	(3.48)	(3.49)	(3.44)
15	0.0270	0.0276	0.0248	0.0284	0.0287
	(0.70)	(0.72)	(0.65)	(0.75)	(0.75)
16	-0.0152	-0.0173	-0.0102	-0.0116	-0.0144
	(-0.47)	(-0.52)	(-0.33)	(-0.38)	(-0.47)
17	0.0204	0.0204	0.00841	-0.0111	-0.00927
	(0.45)	(0.44)	(0.19)	(-0.26)	(-0.22)
Initial LSK intensity	-0.0134	-0.0127	-0.0297	-0.0381	-0.0410
	(-0.32)	(-0.30)	(-0.73)	(-0.96)	(-1.02)
Age in 2006 × LSK intensity					
6	0.0351	0.0349	0.0216	0.0182	0.0331
	(0.46)	(0.45)	(0.28)	(0.25)	(0.44)
7	0.0766	0.0774	0.0658	0.0685	0.0775
	(0.98)	(1.00)	(0.88)	(0.95)	(1.02)
8	-0.0185	-0.0175	-0.0381	-0.0482	-0.0357
	(-0.24)	(-0.23)	(-0.50)	(-0.67)	(-0.47)
9	0.0260	0.0271	0.0303	0.0287	0.0291
	(0.34)	(0.35)	(0.37)	(0.36)	(0.36)
10	-0.00243	-0.00314	-0.0113	-0.00937	-0.000181
	(-0.04)	(-0.06)	(-0.21)	(-0.18)	(-0.00)
11	0.0495	0.0472	0.0339	0.0474	0.0459
	(0.48)	(0.46)	(0.36)	(0.48)	(0.48)
12	-0.0222	-0.0210	0.00625	0.0126	0.0129
	(-0.32)	(-0.31)	(0.10)	(0.20)	(0.20)
13	-0.0457	-0.0491	-0.0621	-0.0692	-0.0565
	(-0.43)	(-0.47)	(-0.63)	(-0.73)	(-0.56)
14	-0.0783	-0.0790	-0.0803	-0.0727	-0.0701
	(-1.03)	(-1.04)	(-1.16)	(-1.07)	(-1.02)
15	-0.0727	-0.0757	-0.0772	-0.0721	-0.0694
	(-1.19)	(-1.23)	(-1.20)	(-1.14)	(-1.08)
16	-0.0401	-0.0397	-0.0418	-0.0380	-0.0367
	(-0.60)	(-0.59)	(-0.66)	(-0.62)	(-0.62)
17	0.0129	0.0104	0.0256	0.0422	0.0295
	(0.17)	(0.14)	(0.37)	(0.65)	(0.48)
Control variables					
Literacy (base line: Cannot read)					
Can read a little		-0.0139	0.00395	0.0178	0.0346
		(-0.11)	(0.04)	(0.17)	(0.36)
Can read		0.0250	0.0271	0.0370	0.0576
		(0.25)	(0.31)	(0.44)	(0.72)

Wealth (base line: low income)					
Middle income			0.0246	0.0198	0.0213
			(0.82)	(0.66)	(0.71)
High income			0.105**	0.0837**	0.0721**
			(3.29)	(3.13)	(2.76)
Language (base line: Other languages)					
English				0.0765	0.0117
				(1.95)	(0.32)
Afrikaans				0.135**	-0.00911
				(3.03)	(-0.24)
Ethnicity (base line: White)					
Black/African					-0.199***
					(-4.92)
Coloured					-0.0515
					(-1.31)
Indian/Asian					-0.0742
					(-1.95)
Constant	0.782***	0.759***	0.721***	0.709***	0.891***
	(30.90)	(7.53)	(8.36)	(8.59)	(11.86)
F-statistic	0.72	0.78	0.81	0.94	0.86
R^2	0.01	0.01	0.03	0.03	0.04
N	4,450	4,447	4,447	4,447	4,447
t statistics in parentheses					
*** p<0.05, ** p<0.01, *** p<0.001**					

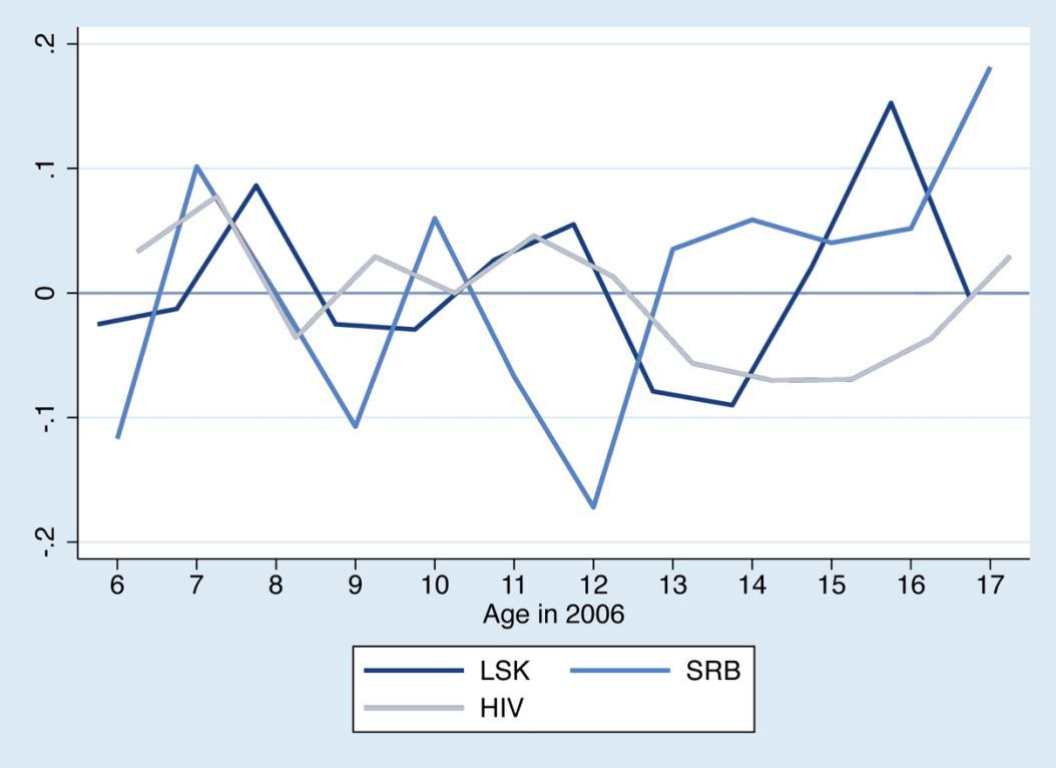
Comment: Column (1) estimates the effect of life skills knowledge without controls. Column (2) adds a variable controlling for literacy, column (3) adds control for wealth, column (4) adds a variable controlling for home language and column (5) adds control for ethnicity. HIV variable is coded so that a value of 1 corresponds to the individual being HIV negative, whilst a value of 0 denotes that the individual is HIV positive. Clustering is made on the basis of municipalities. F-statistics test the hypotheses that the coefficients of interaction between year of birth dummies and the LSK intensity are jointly 0.

Seemingly, the LSE programs has not had any statistically significant effect on HIV prevalence rates. However, this is not surprising considering that there was no statistically confirmed effect on LSK or SRB. Moreover, there is no clear pattern regarding whether the effect increases or decreases with older ages. This is also illustrated in figure 7. Although not statistically significant, figure 7 also provide for a possible comparison of the effect on the different age groups between the model estimating the effect on LSK, SRB and HIV. Similarly, the majority of the variables estimating the average HIV prevalence are not statistically significant. However, the average prevalence of HIV seems to be higher in older ages.

The only statistically significant controls are noted in column 3, 4 and 5. The coefficient for high income takers is statistically significant in all three models (column 3 to 5), thus, indicating that high income takers are 7 to 10 percentage points less likely to have HIV, compared to low income takers. However, the size of the coefficient decreases when including more controls. Additionally, in column 4, the estimate for having Afrikaans as home language is also significant. Consequently, suggesting that individuals with Afrikaans

as home language are 13 percentage points less likely to have a positive HIV status, compared to individuals with a local language as home language. However, the statistically significant result of Afrikaans as home language is lost when controlling for ethnicity (column 5). In the full model (column 5), only the variables measuring high income takers and black/Africans are significant. It is estimated that individuals that are black/Africans are 20 percentage points more likely to have HIV than white people.

Figure 7: Effect of the programs on level of LSK, level of SRB and HIV prevalence rates



Comment: The coefficients of the interaction between year-of-birth dummies indicating age in 2006 and high LSK intensity municipality for each model, LSK, SRB and HIV prevalence rates.

6 Analysis and discussion of results

Based on previous theories arguing that education has a positive impact on health outcomes, this study hypothesises that exposure to life skills education increases an individual's level of life skills knowledge, which in turn results in lower level of sexual risk behaviour and thus, reduced HIV prevalence. The results obtained in the previous sections suggest that there is no statistically significant relationship between exposure to the LSE programs and any of the three outcomes (LSK, SRB and HIV).

Although not statistically confirmed, the effect of the LSE programs on level of life skills knowledge is positive as expected in both the unconditional and short model. However, the results retrieved when estimating the effect in the long model, i.e. on each individual age group, are rather surprising. It was assumed that the effect on level of LSK would decrease with age, since younger individuals have more exposure to life skills education. Seemingly, this is not the case. Instead, the estimated effect is only significant on individuals aged 16 and the effect size fluctuates across all ages. Hence, it is not possible to distinguish whether the effect of the programs increases (or decreases) with increased exposure to the LSE programs. It was also assumed that the average level of LSK should have increased more in municipalities with previously low level of LSK compared to municipalities that were considered as "high initial LSK" municipalities. However, the estimates are not statistically significant, and the generated result suggest the opposite. Instead, the average level of LSK decreases more in "low initial LSK" municipalities. In addition, the average life skills knowledge is 3% to 5% higher among young individuals who attended school in a "high initial LSK" municipality, compared to young individuals who attended school in a "low initial LSK" municipality.

Finally, it was noted that older individuals tend to have on average higher level of LSK compared to younger individuals. Consequently, the results obtained are not aligned with the hypothesised theory, which could be attributed to a number of reasons. One explanation for why older individuals have on average higher knowledge of life skills related issues could be because it is possible to also learn outside of school. Hence, it can be assumed that older individuals have retrieved their knowledge after finishing their mandatory education. Another explanation could be attributed to the simple fact that they are older and perhaps more sexually active. Intuitively, individuals who are sexually active can obtain their knowledge practically in addition to the theoretical knowledge provided in school. Consequently, also implying a higher risk behaviour, however, this will be further discussed below.

Other factors that could impact an individual's ability to obtain high level of LSK are parent's educational attainment, the individual's general health and quality of education (Grossman, 2005:18; Kadakia & Macha, 2017). According to Kadakia and Macha (2017), South Africa has in general higher quality of education compared to other countries in Sub-Saharan Africa. However, it is alarmingly low compared to other countries in the world. In addition, there exist large inequalities between South African schools with regards to available resources. Schools in urban areas are generally wealthier compared to rural schools. Consequently, urban schools are more likely to have higher quality of education and better conditions for learning (Kadakia and Macha, 2017). Thus, it could be argued that the structure of the LSE programs was not adapted and adjusted to match these discrepancies. Instead, a few schools could have benefited more than others from the programs.

Considering that the effect of the programs is statically insignificant on the level of LSK, it is not surprising that the effect of the programs is also statistically insignificant on sexual risk behaviour. Nevertheless, the results of the short and the unconditional model suggest that LSE programs increased the level of sexual risk behaviour by 1% to 3%. It was also hypothesised that the average level of SRB should have decreased more in "low initial LSK" municipalities in 2016, compared to "high initial LSK" municipalities. The findings are in line with the hypothesis, and the average level of SRB is 1% to 7% lower amongst young individuals who attended school in a "low initial LSK" municipality.

Regardless of the overall insignificant result, the long model estimates that the effect of the programs is statistically significant for individuals aged 12 and 17, where the younger group has been negatively impacted whilst the 17-years-old are positively affected. This is not as hypothesised, since the older individuals have had less exposure to the LSE programs. However, since both individuals aged 12 and 17 are statistically insignificant in the model estimating LSK, their significance level in the SRB model should be attributed to reasons that do not directly concern LSK obtained from LSE. Nevertheless, younger individuals are estimated to have on average lower levels of sexual risk behaviour and, building on the argument presented above, older individuals are more likely to be more sexually active, which in turn increases their level of SRB (Wand & Ranjeem, 2012:5-7). Considering that previous studies have established a correlation between level of sexual risk behaviour and the risk of being infected by HIV (NSW Health, 2017:1-3; Wand and Ranjee, 2012:5-7), it is not too surprising that this research also finds that younger individuals have on average lower HIV prevalence, compared to older individuals.

The results also indicate that the average HIV prevalence is 2 percentage points higher amongst young individuals who attended school in a "high initial LSK" municipality, compared to young individuals who attended school in a "low initial LSK" municipality. However, the difference in the decrease is the same in both types of municipalities. Regardless, the estimated effect of the LSE programs on HIV prevalence is statistically insignificant in all models and the direction of the relationship is not as expected in neither the unconditional nor the short models. Thus, it can be concluded that the programs have not had the anticipated effect on HIV prevalence.

Lastly, other general factors that might have an impact on an individual's level of LSK, SRB, and thus, HIV prevalence, are the level of emotional vulnerability, and/or knowing someone who has or has had an STI. More specifically, individuals knowing someone with an STI are more likely to have increased knowledge of the causes and implications of STIs (Onoya et.al, 2014:104-105). Similarly, individuals who have parent's with higher educational attainment tend to also attach higher value to their own education (Grossman and Kaestner, 1997:74).

The included controls provide for some additional information. The individual's home language is most frequently statistically significant, where individuals speaking Afrikaans are estimated to have higher levels of life skills knowledge, lower levels of sexual risk behaviour and lower HIV prevalence, compared to individuals with a local language as home language. Similarly, individuals with English as native language have higher levels of LSK and lower levels of SRB. However, English loses its statistically significant effect in the model estimating the effect on HIV prevalence. Consequently, the findings of this study are in line with previous research stating that individual's with Afrikaans and English tend to have an advantage in obtaining knowledge compared to individuals with a local language as home language (UNICEF, 2016:94).

Additionally, the effect of ethnicity is only statistically significant in the models measuring the impact of the LSE programs on the level of LSK and on HIV prevalence. In the former model, both coloured and black/African individuals are estimated to have lower levels of LSK compared to white individuals. Similarly, black/African individuals have also higher HIV prevalence compared to white individuals. These findings could be attributed to South Africa's history with the apartheid and structural discrimination. According to Oosthuizen (2019), the effects of the apartheid are still present today and non-white individuals are continually discriminated in many aspects of the society. Non-white individuals are also more likely to be low income takers. Hence, when controlling for wealth, it is not surprising that

the estimated effect suggest that high income takers tend to have lower levels of SRB, and HIV prevalence compared to low income takers. Previous research find evidence that low income takers are more likely to engage in multiple sex partner, age-disparate relationships and transactional sex, thus increasing their level of SRB (HSRC, 2019:4-6). Lastly, the level of literacy is statistically insignificant in all models, which is interesting since previous studies have highlighted the importance of literacy in acquiring knowledge (ELINET, 2016:3).

Although previous research studying the effects of the 2005 LSE program found that it increased the knowledge of HIV/AIDS among students in the short-to-medium term (see for instance May et.al. 2004:3; James et.al. 2006:291-292; UNICEF, 2012:84), the results of this thesis suggest that the LSE programs have not been successful in reaching its objectives in the long-term. The HIV prevalence rate remains high and women are still being disproportionally affected by the pandemic. The lack of effect could be due to teachers still being reluctant to discuss sensitive topics as had been argued by May et.al (2004:50). This could also be related to the argument presented by Dupas (2011:433) regarding individuals being responsive to information if the source is credible. If teachers are unengaged or reluctant to teach the subject of life skills, then the information provided might not be considered credible and thus, the students might not be responsive to learn. Consequently, it could be argued that the training of teachers has not been effective in providing them with the appropriate skills and tools to educate on the matter.

Moreover, as was claimed by Prinsloo (2007:164), the failure to acknowledge structural differences between schools in urban and rural areas as well as poorer and richer regions have proven to have had negative impact on the successfulness of the implementation of the LSE programs. Thus, the reason for not yielding the anticipated results could be because South Africa is still struggling with the implementation of the LSE programs, both with regards to the teaching aspect and the overall structure of the programs. However, as stated by the Government of South Africa (2006:46), behaviours are difficult to change. Perhaps the successor of the 2005 and 2011 LSE programs will provide for better outcomes, considering that the subject of life skills will have been more mainstreamed by the time the NSP 2022 has been fully implemented.

To conclude, whether the individual benefits from obtaining life skills knowledge seems to depend on individual characteristics, considering that the individual's wealth, ethnicity and language all tend to matter. It can be assumed that individuals who are either non-white, low income takers and/or have a local language as home language, have lower costs of life skills

knowledge, since they have more to gain from learning. Although this study did not find statistically significant evidence of any of the theorised relationships, there exist previous evidence of a relationship between level of sexual risk behaviour and the risk of being infected by HIV (NSW Health, 2017:1-3; Wand and Ranjee, 2012:5-7). Therefore, by building on the model presented by Dupas (2011), it could be argued that the challenges and costs of the disease are higher than the costs of actually staying in school and being exposed to LSE. Particularly, considering that being infected by HIV has proven to be paralleled with discrimination, stigmatisation and other negative outcomes related to health, social and working life (Spies & Seedat, 2014:6; Santos et.al, 2014:7).

7 Conclusion

The main purpose of this thesis is to study the long-term impact of being exposed to two consecutive LSE programs implemented in South Africa on the level of life skills knowledge, level of sexual risk behaviour and HIV prevalence rates among young women. The main hypothesis tests whether exposure to the LSE programs decreases the level of SRB through increased level of LSK. Subsequently, also decreasing HIV prevalence.

Whether more exposure to LSE actually increases an individual's LSK and decreases its level of SRB and HIV prevalence cannot be convincingly determined. Older individuals are estimated to have on average higher LSK but also higher SRB and HIV prevalence compared to younger individuals, which is not in line with the hypothesised theory. However, this may not be too surprising considering that this research finds no statistically significant evidence of a relationship between exposure to the LSE programs and any of the estimated outcomes. Thus, it can be concluded that the LSE programs have not had the desired or anticipated effect on the level of life skills knowledge, sexual risk behaviour or HIV prevalence. However, as has been discussed above, a number of reasons could be attributed to the unsuccessfulness of the LSE programs. These factors include inadequate implementation, lack of engagement from teachers and low quality of education. Additionally, individual characteristics that could also be of importance are level of sexual activity, parent's education and knowing someone who has or has had an STI.

As this study has been unsuccessful in determining the outcomes of the LSE programs as specified in this thesis, more research on this particular topic should be considered. As a complement to the statistical findings of this study, it would have been beneficial to also conduct interviews on the perceptions of the implementation of the LSE programs among both students and teachers. This would have enabled a clearer understanding of why the LSE programs have not achieved its objectives. Lastly, it would be interesting to also study the effect of being exposed to life skills education among young men, more specifically, to compare the effect between genders, since this could provide for more extensive understanding of the issue. Finally, the HIV pandemic in South Africa remains persistent and severe, thus, it is critical to establish efficient ways to mitigate the problem.

8 Limitations

As many other studies, this research also has some limitations that might impact the results. First, due to data restrictions, only a limited number of variables measuring life skills knowledge were accessible. Thus, the *2016 LSK index* may not cover the whole spectrum of the knowledge and skills that the two LSE programs aimed to develop. Consequently, the LSE programs could have had more positive outcomes on other factors that are also considered as life skills. However, as was argued in the method section, the *2016 LSK index* should be able to measure the most important aspects. In addition, the assumption made with regards to the initial level of life skills knowledge in the included municipalities could also bias the results, particularly with regards to the parallel trend assumption. In order for this research to have obtained more precise estimates, it would have required information on the stepwise implementation of the programs in each school. However, because of limited number of observations and restricted data availability, this was not possible. Instead, the stepwise implementation of the programs in each municipality could have been included. Nonetheless, this information was not available either, thus, using the difference of initial level of LSK across municipalities was the best available option.

Furthermore, as this study excludes municipalities with less than 25 observations, only 34 of South Africa's 52 municipalities were included in the analysis. Thus, the exclusion of the 18 municipalities could have also biased the results. Nevertheless, including these could have posed other implications on the results as some of the municipalities only had 1 observation. Despite the exclusion, all of the nine provinces are represented in the analysis. Thus, it can be assumed that this research has a fairly good national coverage. However, considering that this research is limited to look at South African women and life skills education programs that were specifically developed for the context of South Africa, the generalisability is limited. Consequently, the result obtained here does not necessarily apply to other settings.

Finally, because of data restrictions, the possibility to control for different factors that might impact the effect of the LSE programs was limited. As was discussed, factors such as other sources of knowledge, parent's education, emotional vulnerability, difference in quality of education both between urban and rural areas and across municipalities, may all be of importance. Thus, it would have been beneficial to also control for these in the estimations.

9 Bibliography

9.1 Online sources

Abdi, H. & Valentin, D. (2007) Multiple Correspondence Analysis, [Online] Salkind, N.J., Ed., Encyclopedia of Measurement and Statistics, SAGE Publications, Thousand Oaks, CA, 1-13, available at https://personal.utdallas.edu/~herve/Abdi-MCA2007-pretty.pdf (Accessed: 2020-04-15)

Adewumi, T.M. & Adendorff, M. (2014) Learners' Perceptions of the Life Orientation Curriculum in Fort Beaufort District Eastern Cape in South Africa, [Online] Available at: https://pdfs.semanticscholar.org/aea9/54007564678d89722ed2ef7208e32ba2a853.pdf (Accessed: 2020-05-11)

Aggleton, P., & Clarke, J.D. (2012) Life Skills-Based HIV Education and Education for All, [Online] Education for All Global Monitoring Report, available at: https://unesdoc.unesco.org/ark:/48223/pf0000217866 (Accessed: 2020-05-13)

Magnani, R., MacIntyre, K., Karim, A.M., Brown, L. & Hutchinson, P. (2005) The impact of life skills education on adolescent sexual risk behaviors in KwaZulu-Natal, South Africa, [Online] Journal of Adolescent Health, available at: https://www.academia.edu/3464147/The_impact_of_life_skills_education_on_adolescent_sexual_risk_behaviors_in_KwaZulu-Natal_South_Africa (Accessed: 2020-05-13)

Columbia University Mailman School of Public Health (2019) Difference-in-Difference Estimation, [Online] Columbia University Mailman School of Public Health, available at: https://www.mailman.columbia.edu/research/population-health-methods/difference-difference-estimation (Accessed: 2020-03-28)

Deaton, A. (2013) The Great Escape: Health, Wealth, and the Origins of Inequality, [Online] Princeton: Princeton University Press, available at: http://search.ebscohost.com.ezproxy.ub.gu.se/login.aspx?direct=true&db=nlebk&AN=603306&site=ehost-live (Accessed: 2020-03-30)

Department of Education (2002) Policy: Revised National Curriculum Statement Grades R-9 (Schools), [Online] Available at: https://www.education.gov.za/Portals/0/CD/GET/doc/overview.pdf?ver=2006-11-21-100143-000 (Accessed: 2020-05-02)

Department of Education (2008a) Education for all (EFA) Country report: South Africa, [Online] Available at: https://unesdoc.unesco.org/ark:/48223/pf0000221801 (Accessed: 2020-05-08)

Department of Education (2008b) Education Statistics in South Africa 2006, [Online] Available at: http://www.dhet.gov.za/DHET%20Statistics%20Publication/DoE%20Stats%20at%20a%20Glance%202006.pdf (Accessed: 2020-05-13)

Department of Basic Education (2011a) Curriculum and Assessment Policy Statement Foundation Phase Grades R-3 Life skills, [Online] National Curriculum Statement (NCS), available at: https://www.education.gov.za/Portals/0/CD/National%20Curriculum%20Statements%20and%20Vocational/CAPS%20Life%20Skills%20%20English%20_%20Gr%20R-3%20FS.pdf?ver=2015-01-27-162204-953 (Accessed: 2020-05-02)

Department of Basic Education (2011b) Curriculum and Assessment Policy Statement Foundation Phase Grades 4-6 Life skills, [Online] National Curriculum Statement (NCS), available at: https://www.uj.ac.za/faculties/facultyofeducation/eli/Documents/Life%20Skills%20CAPS%20IP%20GR%204-6%20%20WEB.pdf (Accessed: 2020-05-02)

Department of Basic Education (2019a) HIV and AIDS Life Skills Education Programme, [Online] Government of South Africa, available at https://www.education.gov.za/Programmes/HealthPromotion/HIVandAIDSLifeSkillsEducationProgramme.aspx (Accessed: 2020-03-27)

Department of Basic Education (2019b) School Fees and Exemptions, [Online] Department of Basic Education, available at:

https://www.education.gov.za/Informationfor/ParentsandGuardians/SchoolFees.aspx (Accessed: 2020-05-14)

de Walque, D. (2007) How does the impact of an HIV/AIDS information campaign vary with educational attainment? Evidence from rural Uganda, [Online] The World Bank, Journal of Development Economics 84 (2007) 686–714, doi: https://doi.org/10.1016/j.jdeveco.2006.12.003 (Accessed, 2020-05-14)

Duflo, E. (2001) Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment, [Online] MIT Economics, available at: https://economics.mit.edu/files/726?fbclid=IwAR22vGUw3-k9hVhWjt-BRXxQEkB7noNcc-46m4gjZdV6BDvmHMwPi-gYFVM (Accessed: 2020-03-30)

Dupas, P. (2011). Health behavior in developing countries. [Online] *Annu. Rev. Econ.*, *3*(1), 425–449, avaialble at: https://www.annualreviews.org/doi/pdf/10.1146/annurev-economics-111809-125029 (Accessed: 2020-06-05)

ELINET (2016) European Declaration of the Right to Literacy, [Online] European Literacy Policy Network, available at: http://www.eli-net.eu/fileadmin/ELINET/Redaktion/user_upload/European_Declaration_of_the_Right_to_Literacy2.pdf (Accessed: 2020-05-14)

European Commission (2013) Reducing Early School Leaving: Key Messages and Policy Support, [Online] European Commission, available at: https://ec.europa.eu/education/sites/education/files/early-school-leaving-group2013-report_en.pdf (Accessed: 2020-05-14)

Evans, M., Risher, K., Zungu, N., Shisana, O., Moyo, S., et al. (2016) Age disparate sex and HIV risk for young women from 2002 to 2012 in South Africa, [Online] Journal of the International AIDS Society 19(1): 21310 doi: 10.7448/IAS.19.1.21310 (Accessed: 2020-05-13)

Grossman, M. (2005) Education and Nonmarket Outcomes [Online] Handbook of the Economics of Education, Elsevier, available at: https://www.nber.org/papers/w11582 (Accessed: 2020-05-15)

Government of South Africa (2000) HIV/AIDS/STD Strategic Plan for South Africa 2000-2005, [Online] Government of South Africa, available at: https://www.gov.za/sites/default/files/gcis_document/201409/aidsplan20000.pdf (Accessed: 2020-03-28)

Government of South Africa (2006) HIV/AIDS/STD Strategic Plan for South Africa 2007-2011, [Online] Government of South Africa, available at: http://data.unaids.org/pub/externaldocument/2007/20070604_sa_nsp_final_en.pdf (Accessed: 2020-05-05)

Government of South Africa (2011) National Strategic Plan on HIV, STIs and TB 2012-2016, [Online] Government of South Africa, Available at: https://www.hst.org.za/publications/NonHST%20Publications/hiv-nsp.pdf (Accessed: 2020-05-05)

Government of South Africa (2020) South Africa's People, [Online] The Government of South Africa, available at: https://www.gov.za/about-sa/south-africas-people (Accessed: 2020-05-14)

Hippel, V.P., (2015) Linear vs. Logistic Probability Models: Which is Better, and When?, [Online] Statistical Horizons, available at: https://statisticalhorizons.com/linear-vs-logistic (Accessed: 2020-04-11)

HSRC (2019) South Africa National HIV Prevalence, Incidence Behaviour and Communication Survey, 2017, [Online] The Human Sciences Research Council, available at: http://www.hsrc.ac.za/uploads/pageContent/10779/SABSSM%20V.pdf (Accessed: 2020-05-14)

James, S., Reddy, P., Ruiter, A.C R., McCauley, A. & van den Bornde, B. (2006) The Impact of an HIV and AIDS Life Skills Program on Secondary School, [Online] *Education and*

Prevention, 18(4), 281-94, available at: https://search-proquest-com.ezproxy.ub.gu.se/docview/198002011?accountid=11162 (Accessed: 2020-03-26)

Kadakia, A. & Macha, W. (2017) Education in South Africa [Online] World Education News & Reviews, available at: https://wenr.wes.org/2017/05/education-south-africa (Accessed: 2020-05-10)

Leonard, J. (2020) What is it like to live with HIV?, [Online] Medical News Today, available at: https://www.medicalnewstoday.com/articles/323981 (Accessed: 2020-05-13)

Mabaso, M., (2018) Correlates of Correct HIV Knowledge and Myth Rejection in South Africa: The 2012 National HIV Prevalence, Incidence and Behavior Survey, [Online] J AIDS Clin Res 9: 761, doi: 10.4172/2155-6113.1000761 (Accessed: 2020-05-13)

Mabaso, M., Makola, L., Naidoo, I. Mlangeni, L.L, Jooste, S., & Simbayi, L. (2019) HIV prevalence in South Africa through gender and racial lenses: results from the 2012 population-based national household survey, [Online] Int J Equity Health 18, 167, available at: https://doi.org/10.1186/s12939-019-1055-6 (Accessed: 2020-05-13)

May, J., Manzini, N., Kaufman, C., Rutenberg, N., Hallman, K., Brown, L., Karim M A., Magnani, B., Macintyre, K., Hutchinson, P., Stavros, S., Dallimore, A., & Fensham, R., (2004) Transitions to Adulthood in the Context of AIDS in South Africa: The Impact of Exposure to Life Skills Education on Adolescent Knowledge, Skills, and Behavior, [Online] Horizons Final Report, available at: https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/467/related_materials (Accessed: 2020-03-09)

Motala, S. & Sayeed, Y. (2009) "No Fee" Schools in South Africa. Policy Brief Number 7, [Online] Consortium for Research on Educational Access, Transitions and Equity (CREATE), available at: https://files.eric.ed.gov/fulltext/ED508824.pdf (Accessed: 2020-05-15)

McNeil, J. A. (2019) History of Official Government HIV/AIDS Policy in South Africa. South African History Archives, [Online] South African History Online, available at:

http://www.sahistory.org.za/topic/history-official-government-hivaidspolicy-south-africa (Accessed: 2020-05-12)

Moloi, M. & Strauss, J. (2005) The SACMEQ II Project in South Africa: A Study of Schooling and the Quality of Education, [Online] Harare: SACMEQ, available at: http://www.sacmeq.org/?q=sacmeq-projects/sacmeq-ii/reports (Accessed: 2020-05-09)

National Department of Health, Statistics South Africa (South African Medical Research Council (SAMRC) & ICF (2019), South Africa Demographic and Health Survey 2016, [Online] Pretoria, South Africa, and Rockville, Maryland, USA: NDoH, Stats SA, SAMRC, and ICF, available at: https://dhsprogram.com/pubs/pdf/FR337/FR337.pdf (Accessed: 2020-03-08)

National Department of Health/Statistics South Africa and Macro International (2000) South Africa Demographic and Health Survey 1998, [Online] Pretoria, South Africa, available at: https://dhsprogram.com/pubs/pdf/FR131/FR131.pdf (Accessed: 2020-03-08)

Norrby, E., Malmquist, J,. & Hagberg, L, (n.d) HIV, [Online] Nationalencyklopedin, available at: https://www.ne.se/uppslagsverk/encyklopedi/l%C3%A5ng/hiv (Accessed: 2020-05-12)

NSW Health (2017) HIV infection, [Online] The New South Wales Ministry of Health, available at: https://www.health.nsw.gov.au/Infectious/factsheets/Factsheets/hiv-infection.pdf (Accessed: 2020-05-12)

Onoya, D., Zuma, K., Zungu, N., Shisana, O. & Mehlomakhulu, V. (2014) Determinants of multiple sexual partnerships in South Africa, [Online] Journal of Public Health 37(1): 97–106 doi: 10.1093/pubmed/fdu010 (Accessed: 2020-05-13)

Oosthuizen, M. (2019) Racial inequality and demographic change in South Africa - Impacts on economic development, [Online] UNU-WIDER, available at: https://www.wider.unu.edu/publication/racial-inequality-and-demographic-change-south-africa (Accessed: 2020-05-13)

Prinsloo, E. (2007) Implementation of life orientation programmes in the new curriculum in South African schools: perceptions of principals and life orientation teachers South African Journal of Education, [Online] Available at: https://files.eric.ed.gov/fulltext/EJ1150093.pdf (Accessed: 2020-03-30)

Richter, L., Mabaso, M., Ramjith, J. & Norris, SA. (2015) Early sexual debut: voluntary or coerced? Evidence from longitudinal data in South Africa – the Birth to Twenty Plus study, [Online] *South African Medical Journal, 105(4): 304–307*. doi:10.7196/SAMJ.8925 (Accessed: 2020-05-13)

Santos, d. ML. M., Kruger, P., Mellors, E.S., Woolvardt, G,. & Ryst, d.,v., E,. (2014) An exploratory survey measuring stigma and discrimination experienced by people living with HIV/AIDS in South Africa: the People Living with HIV Stigma Index, [Online] *dos Santos et al. BMC Public Health 2014, 14:80,* available at: http://www.biomedcentral.com/1471-2458/14/80 (Accessed: 2020-05-13)

Simelela, N., Venter, W.D.F., Pillay, Y, & Barron.P (2015) A Political and Social History of HIV in South Africa, [Online] Curr HIV/AIDS Rep 12, 256–261 (2015), available at: https://doi-org.ezproxy.ub.gu.se/10.1007/s11904-015-0259-7 (Accessed: 2020-05-12)

Spaull, N. (2018) Equity: A price too high to pay?, [Online] Available at: https://nicspaull.files.wordpress.com/2018/11/nic-spaull-springer-framing-chapter.pdf (Accessed: 2020-05-13)

Spies, G,. & Seedat, S,. (2014) Depression and resilience in women with HIV and early life stress: does trauma play a mediating role? A cross-sectional study, [Online] BMJ Open 2014;4:e004200. doi:10.1136/bmjopen-2013-004200 (Accessed: 2020-05-13)

Stacey, N. (1998) Social Benefits of Education, [Online] *The ANNALS of the American Academy of Political and Social Science, 559(1), pp. 54–63,* doi: 10.1177/0002716298559001005 (Accessed: 2020-05-16)

Statistics South Africa (2017a) Education Series Volume III Educational Enrolment and Achievement, 2016 , [Online] Statistics South Africa, Available at:

http://www.statssa.gov.za/publications/Report%2092-01-03/Report%2092-01-032016.pdf (Accessed: 2020-03-30)

Statistics South Africa (2017b) Inequality Trends in South Africa A multidimensional diagnostic of inequality, [Online] Available at: http://www.statssa.gov.za/publications/Report-03-10-19/Report-03-10-192017.pdf (Accessed: 2020-05-13)

Stein, N. (2017) Language in Schools [Online] Veriava F, Thom A, Hodgson T (eds) Basic Education Rights Handbook - Education Rights in South Africa, Johannesburg, pp. 205-217, Available at: https://eduinfoafrica.files.wordpress.com/2016/11/basiceducationrightshandbook-complete.pdf (Accessed: 2020-05-13)

UNAIDS (2019) Global HIV & AIDS statistics — 2019 fact sheet, [Online] UNAIDS, available at: https://www.unaids.org/en/resources/fact-sheet (Accessed: 2020-02-08)

UNAIDS (2020) Country – South Africa, [Online] UNAIDS, available at: https://www.unaids.org/en/regionscountries/countries/southafrica (Accessed: 2020-02-08)

UNICEF (2003) Life Skills - Definition of Terms, [Online] United Nations Children's Fund, available at: https://www.unicef.org/lifeskills/index_7308.html (Accessed: 2020-03-27)

UNICEF (2012) Global Evaluation of Life Skills Education Programme, [Online] United Nations Children's Fund, available at: https://www.unicef.org/evaluation/files/USA-2012-011-1_GLSEE.pdf (Accessed: 2020-03-26)

UNICEF (2016) The impact of language policy and practice on children's learning: Evidence from Eastern and Southern Africa 2016, [Online]
Available at: https://www.unicef.org/esaro/UNICEF(2016)LanguageandLearning-SouthAfrica-CaseStudy.pdf?fbclid=IwAR2TvZ9aAemGezrqIKqheLNxvCDnjmBpNjekfeJOFdT8qCCaFJ5-Mats-tE (Accessed: 2020-05-15)

U.S. Department of Labor (c.2014) Findings on the Worst Forms of Child Labor - South Africa, [Online] Available at: https://www.dol.gov/sites/dolgov/files/ILAB/child_labor_reports/tda2014/southafrica.pdf?fbclid=IwAR0JqwfxLtaPAnMZv5gbdehHVguw_nq6ADWhSM2WnVD9Pi1Erk3J9gNQDrc (Accessed: 2020-05-15)

Wand, H. & Ramjee, G. (2012) The relationship between age of coital debut and HIV seroprevalence among women in Durban, South Africa: a cohort study, [Online] BMJ Open 2012;2: e000285. doi:10.1136/ bmjopen-2011-000285, (Accessed: 2020-05-13)

WHO (2019) HIV/AIDS, [Online] World Health Organisations, available at: https://www.who.int/news-room/fact-sheets/detail/hiv-aids (Accessed: 2020-05-13)

WHO (1997) Life Skills Education for Children and Adolescents in Schools, [Online] World Health Organisations, available at: https://apps.who.int/iris/bitstream/handle/10665/63552/WHO_MNH_PSF_93.7A_Rev.2.pdf?sequence=1&isAllowed=y (Accessed: 2020-03-27)

World Education News + Reviews (2017) Education in South Africa, [Online] World Education News + Reviews, available at: http://wenr.wes.org/wp-content/uploads/2017/05/WENR-0517-Country-Profile-South-Africa-1.png (Accessed: 2020-03-31)

Zaidi, J., Grapsa, E., Tanser, F., Newell, M-L. & Bärnighausen, T. (2013) Dramatic increases in HIV prevalence after scale-up ofantiretroviral treatment: a longitudinal population-based HIV surveillance study in rural kwazulu-natal, [Online] AIDS. 2013 September 10; 27(14): 2301–2305. doi:10.1097/QAD.0b013e328362e832 (Accessed: 2020-05-20)

9.2 Printed sources

Aneshensel, C. S., (2013) *Theory-Based Data Analysis for the Social Sciences.* 2nd ed. London: Sage Publication Ltd.

Grossman M & Kaestner R (1997) "Effects of education on health*",* in Behrman J & Stacey N (1997) *The social benefits of education.* Michigan: University of Michigan

Matebesi, S. (2017) *Civil Strife Against local governance: Dynamics of community protests in contemporary South Africa*, Opladen, Berlin & Toronto: Barbara Budrich Publisher

Mehmetoglu, M. & Jakobsen, T. G., (2017). *Applied Statistics Using Stata: A Guide for the Social Sciences*, London: Sage Publication Ltd.

Schmitz, L.C,. Collardey, K.E., Traver K.E., & Larson, D., (2004) Child Labor, Westport. Connecticut: Greenwood Press

10 Appendix

Table A1 presents the included Metropolitan and district municipalities by province and number coding. In total, 34 different municipalities from all nine provinces are included.

Table A1: Included municipalities by province

Code	Municipality	Province
1	City of Cape Town	Western Cape
2	Garden Route District Municipality	Western Cape
3	Chris Hani District Municipality	Eastern Cape
4	Buffalo City Metropolitan Municipality	Eastern Cape
5	Nelson Mandela Bay Metropolitan Municipality	Eastern Cape
6	Pixley ka Seme District Municipality	Northern Cape
7	ZF Mgcawu District Municipality	Northern Cape
8	Frances Baard District Municipality	Northern Cape
9	Lejweleputswa District Municipality	Free State
10	Fezile Dabi District Municipality	Free State
11	Thabo Mofutsanyana District Municipality	Free State
12	Mangaung Metropolitan Municipality	Free State
13	eThekwini Metropolitan Municipality	KwaZulu-Natal
14	uMgungundlovu District Municipality	KwaZulu-Natal
15	Ugu District Municipality	KwaZulu-Natal
16	Umzinyathi District Municipality	KwaZulu-Natal
17	Zululand District Municipality	KwaZulu-Natal
18	Ngaka Modiri Molema District Municipality	North West
19	Bojanala Platinum District Municipality	North West
20	Dr Kenneth Kaunda District Municipality	North West
21	City of Tshwane Metropolitan Municipality	Gauteng
22	City of Johannesburg Metropolitan Municipality	Gauteng
23	City of Ekurhuleni Metropolitan Municipality	Gauteng
24	Gert Sibande District Municipality	Mpumalanga
25	Nkangala District Municipality	Mpumalanga
26	Ehlanzeni District Municipality	Mpumalanga
27	Mopani District Municipality	Limpopo
28	Vhembe District Municipality	Limpopo
29	Capricorn District Municipality	Limpopo
30	Waterberg District Municipality	Limpopo
31	Sekhukhune District Municipality	Limpopo
32	Alfred Nzo District Municipality	Eastern Cape
33	O.R. Tambo District Municipality	Eastern Cape
34	Amajuba District Municipality	KwaZulu-Natal
N	34	9

Table A2 presents the division of municipalities into “high LSK” and “low LSK” municipalities in 1998 and 2016. It also describes the number of observations from each municipality. The intensity of LSK 2016 is mainly included to allow for an overview of the differences and similarities between the initial LSK (1998) municipalities and the current LSK (2016) municipalities.

Table A2: Intensity of initial LSK knowledge by municipality in 1998 & 2016

Code	Intensity of initial LSK 1998	Intensity of LSK 2016	Municipality	1998	1998	2016	2016
				n=0	n=1	n=0	n=1
1	High LSK	High LSK	City of Cape Town	338	225	260	121
2	High LSK	High LSK	Garden Route District Municipality	36	37	43	20
3	Low LSK	Low LSK	Chris Hani District Municipality	224	172	98	50
4	High LSK	Low LSK	Buffalo City Metropolitan Municipality	133	99	134	74
5	High LSK	High LSK	Nelson Mandela Bay Metropolitan Municipality	290	190	102	53
6	Low LSK	High LSK	Pixley ka Seme District Municipality	106	75	71	35
7	Low LSK	High LSK	ZF Mgcawu District Municipality	177	142	109	55
8	Low LSK	High LSK	Frances Baard District Municipality	269	195	138	82
9	Low LSK	High LSK	Lejweleputswa District Municipality	154	116	121	68
10	High LSK	High LSK	Fezile Dabi District Municipality	81	45	64	45
11	Low LSK	Low LSK	Thabo Mofutsanyana District Municipality	159	146	175	106
12	Low LSK	High LSK	Mangaung Metropolitan Municipality	120	86	204	144
13	High LSK	High LSK	eThekwini Metropolitan Municipality	439	333	187	133
14	High LSK	High LSK	uMgungundlovu District Municipality	179	161	76	59
15	Low LSK	Low LSK	Ugu District Municipality	39	44	80	53
16	Low LSK	Low LSK	Umzinyathi District Municipality	71	56	76	51
17	High LSK	Low LSK	Zululand District Municipality	31	28	82	79
18	High LSK	Low LSK	Ngaka Modiri Molema District Municipality	85	104	165	107
19	High LSK	High LSK	Bojanala Platinum District Municipality	84	76	305	215
20	High LSK	High LSK	Dr Kenneth Kaunda District Municipality	116	76	109	74
21	High LSK	High LSK	City of Tshwane Metropolitan Municipality	164	156	152	110
22	Low LSK	High LSK	City of Johannesburg Metropolitan Municipality	210	135	101	75
23	High LSK	High LSK	City of Ekurhuleni Metropolitan Municipality	155	103	150	79
24	High LSK	Low LSK	Gert Sibande District Municipality	202	190	148	100
25	High LSK	Low LSK	Nkangala District Municipality	208	195	227	196
26	Low LSK	High LSK	Ehlanzeni District Municipality	139	134	255	207
27	Low LSK	Low LSK	Mopani District Municipality	172	144	159	97
28	Low LSK	High LSK	Vhembe District Municipality	135	127	191	103
29	Low LSK	Low LSK	Capricorn District Municipality	143	145	85	36
30	High LSK	Low LSK	Waterberg District Municipality	52	43	46	27
31	Low LSK	Low LSK	Sekhukhune District Municipality	29	34	206	139
32	Low LSK	Low LSK	Alfred Nzo District Municipality	242	217	127	86

33	Low LSK	Low LSK	O.R. Tambo District Municipality	326	322	148	112
34	Low LSK	Low LSK	Amajuba District Municipality	62	37	73	60
n				5370	4388	4667	3051
N		34		9758		7507	

Comment: The division is based on the level of life skills knowledge of 18 to 30 years old in 1998 and 2016, respectively. Individuals aged 18 to 30 in are coded as 1, whilst individuals younger than 18 and older than 30 are treated as the baseline group and are coded as 0. Number of observations per group is presented in the table.

Table A3 provide the descriptive statistics of the variables included in the *1998 LSK index*. In panel B, it also includes descriptive statistics of the index used to divide municipalities in "high LSK" and "low LSK" municipalities in 2016.

Table A3: Descriptive statistics, 1998 LSK index variables

	Observations	Mean	Standard deviation	Minimum	Maximum
Panel A: Initial LSK intensity index 98					
Female of interest (aged 18-30 in 1998)	5,257 (11,735[1])	0.448	0.497	0	1
Initial LSK intensity index 98	5,257	-0.001	0.154	-0.425	0.244
Knowledge of contraceptive method	5,257	1.979	0.143	1	2
Knowledge of condom	5,257	1.855	0.352	1	2
Source of condom	5,257	2.080	0.468	1	3
Don't use condom because it's a waste of sperm	5,257	2.325	0.471	1	3
Don't use condom because don't know about condom	5,257	2.294	0.521	1	3
Don't use condom because don't know how to use	5,257	2.304	0.506	1	3
Don't use condom because don't know any source	5,257	2.316	0.486	1	3
Don't use condom because embarrassed to get	5,257	2.322	0.476	1	3
Don't use condom because it will be lost inside	5,257	2.308	0.500	1	3
Don't use condom because low risk of STD	5,257	2.292	0.525	1	3
Don't use condom because it is not cool/trendy	5,257	2.325	0.470	1	3
Don't use condom because partner dislikes	5,257	2.107	0.730	1	3
Don't use condom because religion prohibits	5,257	2.324	0.473	1	3
Protect from HIV by avoid mosquitos	5,257	1.401	0.558	0	2
Protect from HIV by having safe sex	5,257	1.812	0.463	0	2
Protect from HIV by using condom during sex	5,257	1.812	0.463	0	2
Protect from HIV by using clean needles during injections	5,257	1.811	0.463	0	2
Protect from HIV by avoid touching a person with AIDS	5,257	1.505	0.843	0	2
Protect from HIV by having a good diet	5,257	1.615	0.547	0	2
Protect from HIV by avoiding public toilets	5,257	1.619	0.548	0	2
Protect from HIV by avoiding sharing food with person with HIV	5,257	1.662	0.539	0	2
Protect from HIV by avoiding sharing razor blades	5,257	1.797	0.476	0	2
Knowledge of how to protect from HIV	5,257	1.912	0.374	0	2
A healthy-looking person can have HIV	5,257	1.499	0.572	0	2

Panel B: LSK intensity index 16					
Female of interest 2016	3,457 (8,737[1])	0.396	0.489	0	1
LSK intensity index 16	3,457	-0.008	0.191	-0.409	0.423
Knowledge of contraceptive method	3,457	1.999	0.034	1	2
Knowledge of male condom	3,457	1.324	0.468	1	2
Knowledge of female condom	3,457	1.325	0.468	1	2
Drugs to avoid HIV transmission to baby during pregnancy	3,457	1.703	0.677	0	2
Respondent can refuse sex	3,457	0.384	0.748	0	2
Respondent can ask partner to use a condom	3,457	0.397	0.768	0	2
Ever heard of sexually transmitted infection (STI)	3,457	1.712	0.453	1	2

Comment: [1]The total number of observations is 8,737 out of which 5,257 belongs to our group of interest, i.e. females aged 18-30 in 2016.

Table A4 provide for descriptive statistics of the variables included in the main *2016 LSK index* as well as for the variables included in the main *2016 SRB index.*

Table A4: Descriptive statistics, 2016 index variables

	Observations	Mean	Standard deviation	Minimum	Maximum
Panel A: Main Initial LSK index					
Log(LSK) index 2016	8,737	0.933	0.348	2.38e-07	1.493
Knowledge of contraceptive method	8,737	1.997	0.057	1	2
Knowledge of male condom	8,737	1.371	0.483	1	2
Knowledge of female condom	8,737	1.339	0.473	1	2
Drugs to avoid HIV transmission to baby during pregnancy	8,737	1.694	0.682	0	2
Respondent can refuse sex	8,737	0.573	0.853	0	2
Respondent can ask partner to use a condom	8,737	0.590	0.871	0	2
Ever heard of sexually transmitted infection (STI)	8,737	1.695	0.460	1	2
Panel B: SRB index					
Log(SRB) index 2016	8,737	0.474	0.450	0	1.449
Current use of contraceptive method	8,737	1.471	0.499	1	2
Condom used during last sex with most recent partner	8,737	1.802	0.787	1	3
Ever had sex	8,737	1.091	0.288	1	2
Number of sex partners, including spouse, in last 12 months	8,737	1.230	0.421	1	2
Had STI in last 12 months	8,737	1.947	0.223	1	2
Ever tested HIV	8,737	1.860	0.347	1	2
Relationship to most recent sexual partner	8,737	2.213	0.449	1	3

Table A5: Correlation matrix

	LSK	SRB	HIV status	Young	LSK intensity	Age 6	Age 7	Age 8	Age 9	Age 10	Age 11	Age 12	Age 13	Age 14	Age 15	Age 16	Age 17	Wealth	Language	Literacy	Ethnicity
LSK	1.0000																				
SRB	-0.3252*	1.0000																			
HIV	-0.0034	0.0352*	1.0000																		
Young	-0.3721*	0.2360*	0.0707*	1.0000																	
LSK	0.0369*	-0.0212	-0.0399*	-0.0171	1.0000																
Age 6	-0.1273*	0.2935*	0.0316*	0.1701*	-0.0272	1.0000															
Age 7	-0.1574*	0.2665*	0.0357*	0.2134*	-0.0468*	-0.0339*	1.0000														
Age 8	-0.1300*	0.2019*	0.0311*	0.2050*	-0.0018	-0.0326*	-0.0409*	1.0000													
Age 9	-0.1187*	0.1042*	0.0161	0.2160*	-0.0352*	-0.0344*	-0.0431*	-0.0414*	1.0000												
Age 10	-0.1038*	0.0543*	0.0302*	0.2335*	-0.0356*	-0.0371*	-0.0466*	-0.0447*	-0.0471*	1.0000											
Age 11	-0.0847*	-0.0228	0.0073	0.2092*	-0.0037	-0.0333*	-0.0418*	-0.0401*	-0.0423*	-0.0457*	1.0000										
Age 12	-0.0845*	-0.0448*	0.0104	0.2286*	0.0220	-0.0364*	-0.0456*	-0.0438*	-0.0462*	-0.0499*	-0.0447*	1.0000									
Age 13	-0.0501*	-0.0163	0.0129	0.2150*	-0.0098	-0.0342*	-0.0429*	-0.0412*	-0.0434*	-0.0469*	-0.0421*	-0.0459*	1.0000								
Age 14	-0.0412*	-0.0404*	0.0030	0.2354*	0.0372*	-0.0374*	-0.0470*	-0.0451*	-0.0475*	-0.0514*	-0.0460*	-0.0503*	-0.0473*	1.0000							
Age 15	-0.0253	-0.0374*	-0.0024	0.2216*	0.0219	-0.0352*	-0.0442*	-0.0425*	-0.0448*	-0.0484*	-0.0434*	-0.0474*	-0.0445*	-0.0488*	1.0000						
Age 16	-0.0093	-0.0599*	-0.0140	0.1928*	0.0365*	-0.0307*	-0.0385*	-0.0370*	-0.0389*	-0.0421*	-0.0377*	-0.0412*	-0.0388*	-0.0424*	-0.0400*	1.0000					
Age 17	0.0141	-0.0376*	0.0147	0.1643*	0.0020	-0.0261	-0.0328*	-0.0315*	-0.0332*	-0.0359*	-0.0321*	-0.0351*	-0.0330*	-0.0362*	-0.0340*	-0.0296*	1.0000				
Literacy	0.1185*	0.0188	0.0642*	-0.0514*	0.1384*	0.0109	-0.0577*	-0.0001	-0.0344*	0.0026	-0.0049	-0.0173	-0.0002	-0.0360*	0.0164	0.0218	0.0358*	1.0000			
Wealth	-0.1183*	-0.0558*	-0.0855*	0.0323*	-0.1091*	0.0059	0.0443*	-0.0142	0.0087	0.0007	0.0148	0.0264	0.0219	-0.0036	0.0081	-0.0160	-0.0365*	-0.2548*	1.0000		
Language	0.0078	-0.0013	0.0452*	0.0371*	0.0247*	0.0150	-0.0107	-0.0010	0.0141	0.0113	0.0114	0.0125	-0.0268	0.0152	0.0327*	0.0381*	0.0128	0.1740*	-0.0470*	1.0000	
Ethnicity	0.1493*	0.0448*	0.1233*	-0.0436*	0.0434*	0.0115	-0.0390*	0.0116	-0.0115	-0.0061	-0.0256	-0.0259	-0.0221	-0.0175	-0.0146	-0.0119	0.0689*	0.3022*	-0.5943*	0.0368*	1.0000

*Comment: It is measured by Pearson's R. 1 indicates perfect correlation and 0 no correlation. *p<0.05, **p<0.01, ***p<0.001*
The cohort of exposure is described by variables for age 6 to age 17

As can be seen in figure A1, all models have issues with non-normal distribution of residuals However, this is not of big concern, since non-normal distributed errors are mainly problematic in models with small samples (Mehmetoglu and Jakobsen, 2017:151-153).

Figure A1: Descriptive statistics, normality of residuals

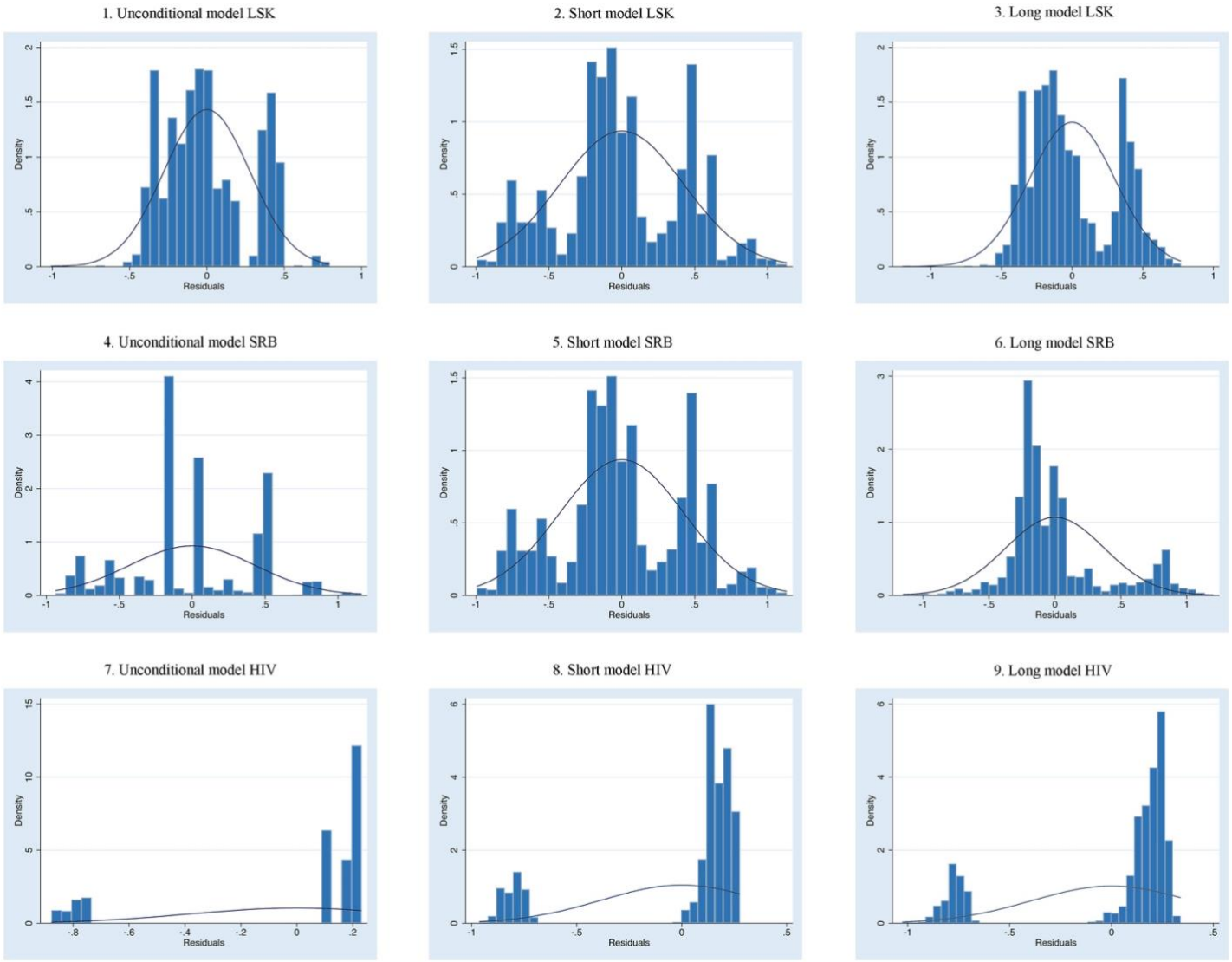


Figure A2 displays the results from the Cook's distance estimate. A few observations in each of the nine estimated models exert a bit more influence than others. However, Mehmetoglu and Jakobsen (ibid) suggest a cut-off value of 1 and none of the observations in any of the models exceeds a Cook's distance value of 1. Consequently, it is concluded that influential observations should not constitute a problem.

Figure A2: Descriprive statistics, presence of influential observations

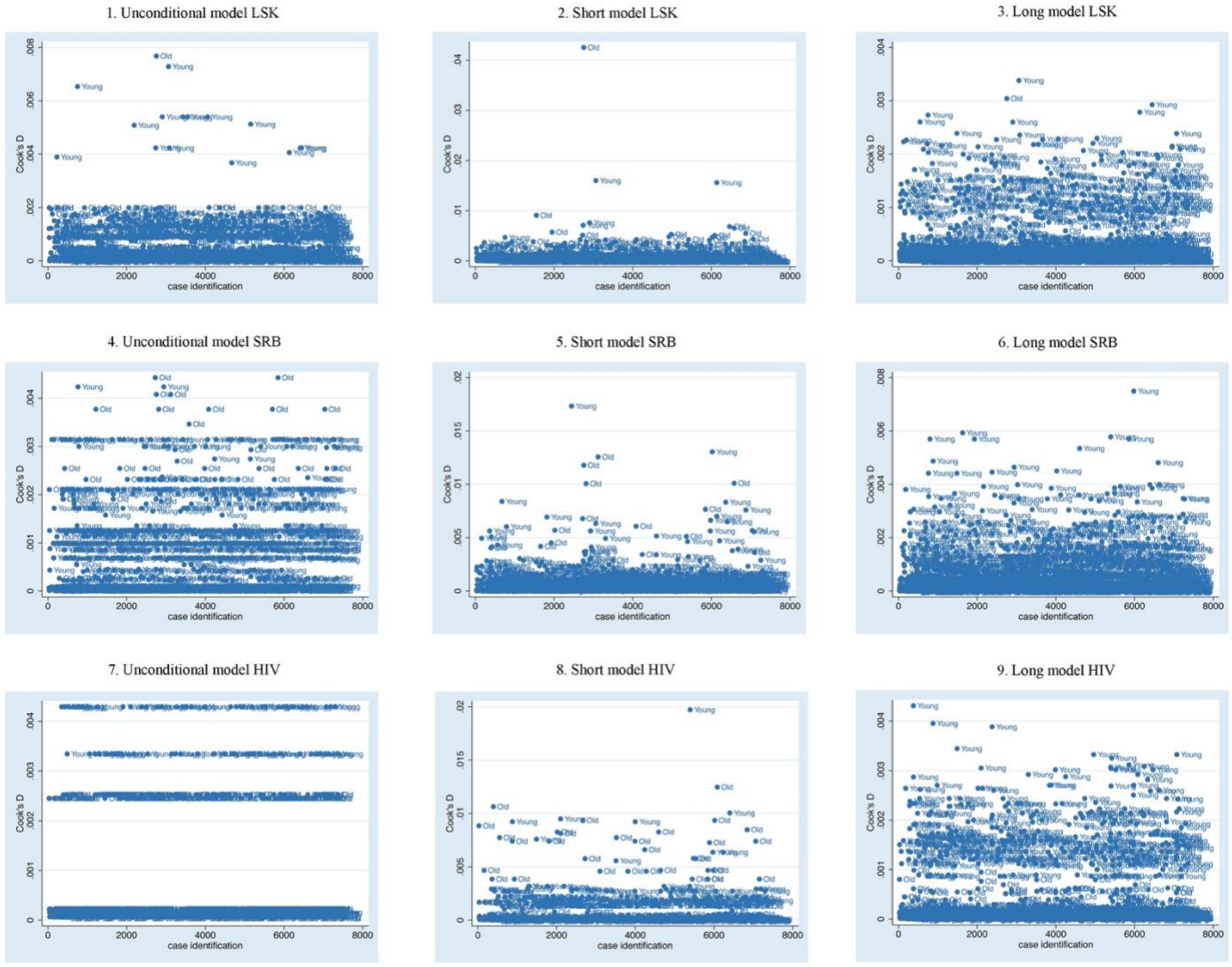


Table A6 presents the mean Variance Inflation Factor (VIF)-value from the test assessing multicollinearity in each of the nine models. Mehmetoglu and Jakobsen (2017: 146-147), suggests a cut-off value of 5 and as can be seen, all of the models have a VIF-value below 5. Thus, it can be concluded that none of the models have problems with multicollinearity.

Table A6: Diagnostic test, multicollinearity by mean VIF-value

Unconditional model: LSK	Short model: LSK	Long model: LSK
2.03	1.62	1.94
Unconditional model: SRB	**Short model: SRB**	**Long model: SRB**
2.03	1.62	1.94
Unconditional model: HIV	**Short model: HIV**	**Long model: HIV**
2.03	1.62	1.94