UNIVERSITY OF GOTHENBURG
SCHOOL OF BUSINESS, ECONOMICS AND LAW

Master Degree Project in Economics

Spend Now, Worry later

The impact of positive income shocks on HIV incidence in Northern Malawi

Dante Anderson and Frida Larsson

Supervisor: Annika Lindskog
Master Degree Project No. 2014:62
Graduate School

UNIVERSITY OF GOTHENBURG
SCHOOL OF BUSINESS, ECONOMICS AND LAW

Spend Now, Worry Later

The Impact of Positive Income Shocks on HIV Incidence in Northern Malawi

Dante Anderson
(890325-5936)

Frida Larsson
(870424-4980)

Supervisor: Annika Lindskog

5th of June 2014

Abstract:
In this study we examine the relationship between the establishment of the Kayelekera Uranium Mine in northern Malawi and HIV/AIDS incidence. The establishment of mining industries is on the rise in this region and the subsequent health impacts are a major concern to governments and the general public. To analyse this topic, we use both a qualitative and a quantitative approach. We find that HIV incidence peaked during 2009 and that this to a large extent is explained by higher income levels in the district. Higher incomes were shown to affect HIV prevalence through increased consumption of alcohol and subsequent risky sexual behaviour. This paper contributes to the literature of how large foreign direct investments and positive income shocks affect HIV/AIDS incidence and risky sexual behaviour in Sub-Saharan Africa.

Acknowledgements

We would like to thank the Elof Hansson Foundation for providing us with the scholarship that enabled our field study in Malawi. The opportunity to travel to Malawi gave us useful insights and enriched our understanding of the local culture and how the mine establishment has affected individuals in the region.

We would also like to extend our thanks to Gillian Swaby and Greg Walker at Paladin Africa Ltd. as well as Jim and Robyn Nottingham, without whom this paper, and our stay in Karonga, would not have been as fulfilling or pleasant. Many thanks, also to all the interviewees who took the time to sit down and share their views of the situation with us.

Lastly, we would like to thank our supervisor Annika Lindskog for her guidance when writing this thesis. Thank you for the helpful comments, suggestions and for always being available for questions.

Table of Contents

1 Introduction

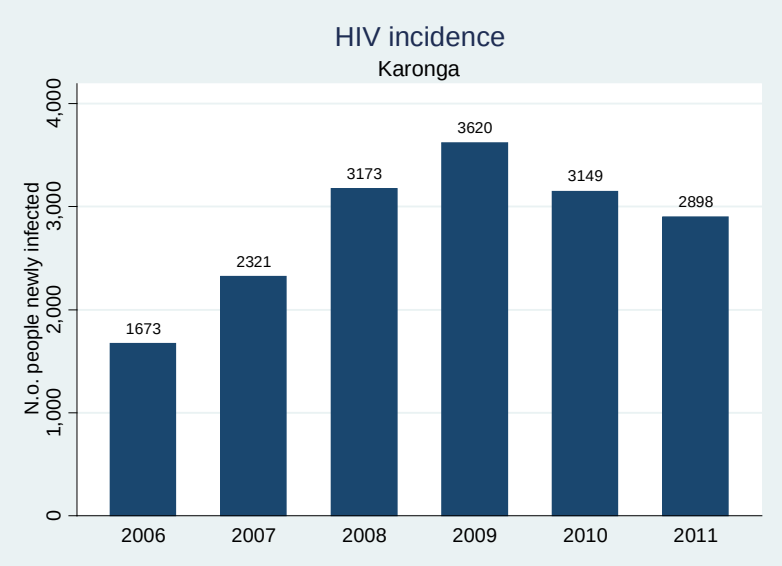
In June 2007, Paladin Africa Ltd. began construction on what would become the Kayelekera Uranium Mine in the Karonga district of northern Malawi. This is to date the largest foreign direct investment in Malawi, a country that suffers from massive poverty and high prevalence of HIV/AIDS. In this paper we will investigate how the establishment of the mine industry has affected HIV incidence in the region, as well as changes in several indicators of risky sexual behaviour and income inequality as possible mechanisms driving the HIV/AIDS epidemic.

The opening of a large mining operation in a Sub-Saharan African country, such as Malawi, is bound to be accompanied by significant changes in the everyday life of people living in the nearby community. It will have significant impacts on, for example, local infrastructure, communications and environment. The increased economic activity ought to have positive impact on local employment, as supported by Kotsadam and Tolonen (2014), which in turn will aid in raising incomes for several individuals in the region. For our thesis the most interesting impact is how the increased economic activity in the area affects sexual behaviour and income inequality. We consider the establishment of the Kayelekera Uranium Mine as a natural experiment, as it is a large exogenous event that affects the lives of most people in the district.

Increased income inequality brought on by the establishment of mining operations is a big issue in Sub-Saharan African communities. Income inequality can have severely detrimental effects on the well-being of the inhabitants of the area, specifically, for the purposes of our thesis, with regards to the spread of HIV and AIDS. Increased income inequality can affect HIV prevalence rates both through increased income as well as through decreased incomes or higher prices within the local commodity market. Higher income has been proven as a driver of risky sexual behaviour in Sub-Saharan Africa, particularly among men, whereas lower incomes can force women into transactional sex or sex-for-transfers.

Figure 1 shows the number of new HIV infections in the Karonga district between the years 2006 and 2011. This data, collected from the Ministry of Health of Malawi, shows an increase

Figure 1: The number of new HIV-infections in the District of Karonga, by year



in infections with a peak in year 2009, after which the numbers start to decline. This peak coincides with the opening of the mine and the final stages of construction.

We pose two research questions:

- What has happened to HIV/AIDS incidence in Karonga since the establishment of the Kayelekera Uranium Mine?
- What has happened to the different mechanisms driving the HIV epidemic in Karonga, e.g. income inequality and risky sexual behaviour?

In order to answer these questions, we will combine qualitative and quantitative methods. The former consists of a field study performed in Malawi during March 2014. The field study is made up of 18 interviews with key informants, such as people employed at the mine, health workers and NGOs. The results from these interviews show that as people receive relatively high wages from the mine, alcohol consumption increases. Individuals with high incomes are expected to treat friends and neighbours to alcohol, which spreads the alcohol consumption to more people than those who are employed at the mine. When these people are inebriated they seem to disregard what they know about safe sex and engage in risky sexual behaviour, such as unprotected or extramarital sex.

The quantitative part of our research includes two datasets. The first from the Malawi Demographic and Health Surveys (MDHS/DHS) is used to analyse changes in HIV/AIDS and

risky sexual behaviour. The second dataset from the Living Standards Measurements Studies (LSMS) is used to analyze income and wealth inequalities. The data stems from two time periods, 2004 and 2010, meaning that they represent both how it was before the mine was opened and after. This will make it possible to use difference in differences-estimation in order to control for time- and district-invariant omitted variables and find the true effect of the mine establishment in Karonga. The results from the analysis of the DHS-data show that there is no coherent change in the data with regards to risky sexual behaviour as a result of the establishment of the Kayelekera Uranium Mine. The likelihood of women being virgins until marriage, the number of women having children during their teens and the likelihood that a woman's partner is dominant all increased. However, this does not imply a clear change in the pattern of risky sexual behaviour, or a consistent impact on HIV and AIDS.

The results from the LSMS-data analysis indicate that income inequality has increased in Karonga as a result of the mine establishment. The high wages has increased the number of relatively rich people in the area. The increased income levels have put an upward pressure on prices, resulting in inflation, which have made people that are not receiving a share of the high mining salaries, relatively poorer. These findings are in line with the results from the qualitative study, which indicates that the wealth has not been equally spread.

This study adds to the existing literature on positive income shocks, stemming from large exogenous events, in very poor areas and their effect on HIV/AIDS incidence and wealth inequality. Similarly it also adds to the literature on the economic and health impacts of the establishment of mines in Sub-Saharan Africa. This paper also contributes to the existing literature on the relationship between income inequality and HIV/AIDS in Malawi.

This paper is organized as follows. Section 2 presents background information about Malawi, the spread of HIV/AIDS and the Kayelekera Uranium Mine. Section 3 presents the existing economic theory behind sexual behaviour and how it is connected to the spread of HIV/AIDS. In section 4 we examine the existing empirical findings of the potential mechanisms driving the HIV/AIDS epidemic. Section 5 describes both the qualitative and the quantitative method used and section 6 introduces the two datasets used MDHS and LSMS. In section 7 we present the results from the qualitative and quantitative research and in section 8 we perform robustness checks on the quantitative findings. Finally, in section 9, we conclude and discuss the findings from the qualitative and quantitative research.

2 Background

In the first section the country of Malawi and the HIV/AIDS situation will be addressed. The second part will present the history and current status of the Kayelekera Uranium Mine.

2.1 Malawi and HIV

The republic of Malawi is one of the least-developed countries in the world, with an economy highly dependent on agriculture. It has a population of 14 million people living on a landmass of 118,000 square kilometres, making it one of the most densely populated nations in the world. The largely rural population grows sugarcane, cotton, corn, and potatoes, and holds cattle and goats. The main export goods include tobacco, tea and sugar. The country is heavily dependent on foreign aid to meet development needs, and suffers from food insecurity and other similar problems caused by the societal underdevelopment (Wielga, 2010).

Malawi is suffering from a high prevalence of HIV/AIDS, according to estimates from the MDHS (2010), the rate is approximately 11 percent among adults aged 15 to 49. This is above the general infection rate in Sub-Saharan Africa, where the prevalence rate is 7.2 percent among the adult population. It is however, below some of the worst infected countries, such as Swaziland with a prevalence of 38.8 percent and Namibia where 21.3 percent of all adults are infected (Canning, 2006). The HIV rate in Malawi has been relatively constant since the mid 1990s, indicating that the disease has settled into a steady state.

The two most prevalent infection channels are heterosexual sex and mother to child transmission. Similar to other Sub-Saharan countries, the females are the ones worst affected by the epidemic. According to the MDHS 2010, more than half of the infected individuals are females. This is different from the HIV infections in the Western world were two thirds of the infected were males (Greenwood et al., 2013). The males are overrepresented and the major transition channels are through drug use and homosexual sex. In Sub-Saharan Africa, by comparison, HIV/AIDS is a disease that disproportionally affects married, divorced and widowed women. There are large differences between the genders and women get infected at younger ages than men. Most individuals know about HIV and how to protect themselves according to MDHS 2010. Condoms are used by less than half of the population and Malawian women have their sexual debut earlier then Malawian men. This is not surprising since large age gaps are common in sexual relationships. There is still stigma and

discrimination against individuals that are HIV positive and these attitudes might negatively affect people's willingness to go and test themselves for HIV.

2.2 Kayelekera Uranium Mine

In the early 1980`s the Central Electric Generation Board of Great Britain (CEBG) discovered a high grade sandstone uranium deposit in Kayerekera village in Malawi. The company worked on the project over an 8-year period and concluded that the project would be unprofitable due to low uranium prices and the high cost of extraction. Therefore, the project was abandoned and Paladin (Africa) Limited, a Malawian company owned by the Paladin Energy Group of Companies of Australia acquired a 90% share in the deposit through a joint venture with Balmain Resources Pty Ltd, the company that held the exploration rights over the area at the time. In July 2005, Paladin Africa Ltd. acquired the remaining shares in the company, but in 2009 the Malawian government was given 15% of the ownership (Paladin Energy Ltd, u.d.).

The Kayelekera Uranium Mine is located in northern part of Malawi, 52 kilometres west of the provincial town of Karonga. The mine is owned and operated by Paladin Africa that has invested US$ 500 million in the mine, making it the largest foreign direct investment project in the country so far. A corporate development agreement was signed by Paladin and the Malawian government on February 22nd 2007. In the agreement, Paladin pledged to spend $10 million on social development projects in the region (Paladin Energy Ltd, u.d.).

Open pit mining is used since the minerals are located relatively near the surface. The low-grade uranium mined contains low levels of uranium oxides. Hence, extensive measures must be undertaken to extract the metal from its ore. The uranium ore is crushed and rendered through an advanced operational process and the resulting product is yellow cake uranium. This product is then transported via trucks to the Walvis Bay in Namibia, and then shipped to the Canadian port Sorel in Montreal. The yellow cake is then refined and converted into pure uranium hexafluoride in Canada (Paladin Energy Ltd., 2012).

During the construction phase in 2008, the number of employees peaked with around 2,000 persons, at least 75 percent being Malawian workers. The mine officially opened in April, 2009. During the extraction phase the labour force originally consisted of 1,100 workers. Due to a number of factors, however, including natural layoffs, deaths and workers quitting, the

workforce gradually decreased over time. The mine needs around 350 workers to operate, which indicates that Paladin Africa has a higher employment rate than needed. This is, according to the General Manager of Paladin Africa, because the company is taking social responsibility, as the employees have no outside option of work. After the natural catastrophe in Fukushima, the uranium spot price dropped sharply and the demand for uranium on the world market declined. As a result, the company decided to put the Kayelekera Uranium Mine on care and maintenance in February 2014. It is expected that once the uranium price recovers, the mining and milling operations will be re-started. Before the company decided to put the mine on care and maintenance, 724 Malawians were employed together with 40 expatriates. These rates will decline as time passes, to around 120 and 26 respectively.

3 Theory

The economic literature presents few formal models explaining the relationship between risky sexual behaviour[1] and HIV/AIDS. Most of the existing work has focused on only one dimension of risky sexual behaviour and does not take into account inter-temporal considerations of an individual over his or her lifetime. For example, Kremer and Morcom (1998) built a model in which the number of partners is decided by a rational individual considering his or her lifetime likelihood of contracting HIV.

The only other formal model that includes men and women is a matching model of marital search conducted by Magruder (2011). The individuals use backwards induction to solve the matching problem and the decision to enter a sexual relationship is not based on the prevalence rate of HIV/AIDS in society. Instead couples have sex, without any decision regarding the use of condoms, to explore whether or not they are a good match for each other. This model implies that the infection risk is more or less constant for the uninfected single, while it declines exponentially with marriage. Hence, uninfected individuals may self-select into the safety of marriage, while infected persons have less to gain from a marriage and might therefore engage in more risky actions.

Morris and Kretzschmar (1997) apply a fully stochastic model to explain the high HIV/AIDS prevalence in Sub-Saharan Africa, by simulating the impact of concurrent relationships.

[1] Risky sexual behaviour indicates that an individual has a higher risk of attracting different kinds of sexual transmitted infections, such as Chlamydia and AIDS. This can for example be that individuals have unprotected sex with multiple partners or interact with the high risk groups like sex workers or fishermen.

Sequential monogamy is compared to a rising level of concurrency, including ten different scenarios, holding all other features of the infection process constant. The authors show that the concurrent partnerships have an important impact on both the size and speed of the spread of the infection in Sub-Saharan Africa. These theories are not likely to be affected by the establishment of a mine, as they mainly deal in set attitudes and norms that already exist.

Further, the authors Greenwood et al. (2013) follow Kremer and Morcom (1998) and apply a choice-theoretical general equilibrium search model to analyze the HIV epidemic in Malawi. The individuals make rational choices about whether to engage in risky sex, based on the odds of contracting HIV. The HIV-positive individuals have less to lose and are more likely to engage in risky sexual behaviour. If there are more men interested in short-term risky sexual relationships, monetary transfers can be used to attract women to that specific market to reach equilibrium. The model is adjusted to fit Malawi and gives mixed results. Some policies, like the promotion of condoms can result in higher HIV prevalence, and the authors emphasize the need for further research on the behavioural change among affected individuals. With regards to a mine establishment, it is possible that it brings an inflow of individuals with preferences towards risky sexual behaviour. This, combined with the economic activity of the mine, can increase the willingness to pay for sex for the in-migrants and thereby increase risky sexual behaviour.

Oster (2012) discusses the connection between high HIV prevalence and changes in economic activity in Sub-Saharan Africa. She builds a simple model, in order to examine this relationship including two main features. The first part describes the relationship between sexual behaviour and economic activity. A country consists of high-risk and low-risk individuals, where the former engages in disproportionately more risky sexual activities than the latter. The high-risk individuals can, for example, be truckers or miners and an increase in exports would increase the number of high-risk individuals, as these types of jobs are associated with higher exports. This relationship can also be seen through income; the high-risk individuals are the ones with relatively higher incomes and these incomes will also increase with higher exports. The second important feature is the relationship between high-risk individuals and the spread of HIV. In Oster`s model the low-risk individuals search for a partner among the general population, while the high-risk group also interact with an outside population i.e. sex workers. The outside population have, on average, higher HIV rates than the general population. These two parts are taken together to estimate the relationship

between infection rate and log exports. She concludes that the number of new infections increases with exports in Sub-Saharan Africa, implying that an increased production in the mining sector could increase HIV incidence indirectly through exports.

These pieces of literature mainly discuss the connection between sexual behaviour and the spread of HIV/AIDS. There are however, other theories that examine the association between income inequality, risky sexual behaviour and their link to HIV. Some authors argue that higher income enables individuals to engage in more risky sexual behaviour since they, for example, can afford to buy unprotected transactional sex. Others argue that it is poverty and not income inequality that drives the epidemic, while some put forth the opposite argument. The link between alcohol and risky sexual behaviour is also of great importance. Hence, there are a number of potential mechanisms driving the HIV epidemic and while the impact of a mine in Sub-Saharan Africa is not well analysed in the formal literature, there are several empirical papers testing these potential transition channels.

4 Empirical findings

Kotsadam and Tolonen (2014) put forth some empirical evidence that mine establishments have the ability to create job openings and cash-earning opportunities outside the agricultural sector. In their analysis of 109 openings and 84 closings of industrial mines in Africa, they find that it is mostly males who switch from working in the agricultural sector to skilled manual labour positions. Thus, mines can increase income for people living in the nearby area, albeit favouring men in a disproportionate manner.

High income and relative wealth can have significant impacts on an individual's probability of contracting HIV. Oster (2009) finds that high income tends to make an individual engage in risky sexual behaviour, such as having multiple sexual partners, frequently having unprotected sex and having sex while intoxicated. Individuals who face lower non-HIV mortality risks, and richer people in general, tend to be more likely to change their behaviour in response to the HIV/AIDS-epidemic. These people are more forward-thinking than others and stand to lose comparatively more if they were to contract HIV. Gertler et al. (2005) find that it is more costly to engage in unprotected transactional sex than protected, which implies that it is more probable that rich individuals are the ones who engage in unprotected transactional sex. Parkhurst (2010) argues that how wealth status affects HIV prevalence rates

in Sub-Saharan Africa depends on underlying structures within the specific countries, and fails to find a common pattern within the continent. However, his research specifically focusing on Malawi, suggests that among males the highest prevalence is found in the highest earnings quintile.

What then are the effects of these income increases, stemming from the expansion of the mining sector, on HIV/AIDS prevalence in Sub-Saharan Africa? Wilson (2010) looks at the effect of a 400% increase in the price of copper on HIV/AIDS prevalence rates in Zambia, a country highly dependent on copper production. This copper boom induced reductions in the risky sexual behaviour that drives the HIV/AIDS epidemic in the area closest to the mine-site. The author saw decreases in transactional sex, multiple partnerships, alcohol use and sex, coital frequency, pregnancy rates and marital rates, with the most change occurring among the young and in-migrants. On the contrary, Corno and de Walque (2012) found that migrants coming to work at mines tend to have higher HIV prevalence rates than locally employed mine workers, meaning that it is more likely that it is the migration, rather than the mine itself, that is driving the higher prevalence. Oster (2012) finds that the number of new HIV infections increase when exports increase. As most mining production in Sub-Saharan Africa is aimed at being exported, her results suggest that increased production in the mining sector ought to increase HIV incidence indirectly through exports.

There are programs aimed at using economic compensation to incentivize safe sex and reductions in risky sexual behaviour. Baird et al. (2012) find in their randomized control trial that school girls who received unconditional or conditional cash transfers have lower levels of HIV prevalence after 18-months. This is in line with the findings by de Walque et al. (2011), who shows that cash transfers, conditional on testing negative for a set of sexually transmitted infections and HIV, reduces risky sexual behaviour and the risk of contracting these infections among adults in rural Tanzania . Kohler and Thornton (2011), however, found no difference in sexual behaviour among the participating men and women in rural Malawi, when offering conditional cash transfers to those who were able to maintain their HIV status for one year. What they did find was that one week after the program had been completed and payments given out, males were 9 percentage points more likely to engage in risky sexual behaviour, whereas females were 6.7 percentage points less likely to do so. This implies that the impact of positive income shocks have increasing effects on males' HIV prevalence rates and decreasing on females'.

While positive income shocks can increase the probability of contracting HIV, the authors Burke et al. (2011) finds that negative income shocks also increases the risk of getting infected. They use sustained periods of drought as a proxy for negative income shocks, as these are likely to affect incomes negatively in societies dependent on agricultural productivity. They find large significant effects on HIV prevalence rates, specifically for people in rural areas.

Wielga et al., (2010) examine how the Kayelekera mine establishment affected human rights and show that the increase in local employment put an upward pressure on price levels. The authors also discuss how the inflation affects the poorest individuals the worst. The people who do not benefit from the income increases created by the mining sector will become relatively poorer as their cost of living will increase. Some people getting richer and some relatively poorer means that income inequality increases. Increased income inequality can itself have detrimental impacts on HIV prevalence rates. Holmqvist (2009) performs a meta-study as well as a cross-country regression and finds that there is a clear positive relationship between HIV and income inequality in Sub-Saharan Africa. Durevall and Lindskog (2011) confirm that the relationship between income inequality and HIV also exist within countries, in their analysis of young women in Malawi. The author Ashley M. Fox (2010) presents a literature review of the growing evidence that wealthier countries and wealthier individuals within countries in Sub-Saharan Africa have a higher risk of attracting HIV. The author contradicts the common knowledge that poverty is the main driver of the epidemic and states that growth that comes on the expense of equity might even boost the spread of HIV.

The relationship between income inequality and the spread of HIV/AIDS has also been tested more carefully by Swindler and Watkins (2007). The authors examine the role of patron-client ties, what Western observers commonly refer to as “transactional sex”. They show that transactional sex is not only about poor women, but also about wealthy men that are expected to take many sexual partners to spread their wealth. It can also be that women, who are not extremely poor but who want to reach a higher economic independency, or simply enjoy some consumption goods like soap or lotions, might engage in these sex-for-transfers. Their results are in line with the findings by Michelle Poulin (2007), who examines the relationship between sex and money in the Balaka region of Malawi. Her deep-interviews show that giving and receiving gifts is the social norm in these societies. The males give money or other

gifts to show affection and appreciation, and a sudden decline in the cash flow could result in a breakup. For females, receiving money means maintaining respect and avoiding ridicule. Tawfik and Watkins (2006) obtain similar results when looking at different interpretations of gender and HIV transmission. The authors show the same pattern, that women do not only sell their bodies to attain basic needs, but that transactional sex is also motivated by the ability to purchase consumer goods and to fulfil sexual desires. There is also a study by Nancy Luke (2005), which uses survey data from men in Kisumu, Kenya, to show that sugar daddies play a significant role in the spread of HIV/AIDS. Rich men are more likely to engage in transfer-for-sex-relationships, albeit being more likely to adhere to the message of condom use (Luke, 2005).

Neo K. Morojele et al. (2006) show that strong link exists between alcohol consumption and risky sexual behaviour in Gauteng, South Africa. Similar results have been found in other southern African countries like Zimbabwe, where alcohol use and risky sexual behaviour are the norm rather than the exception when it comes to the spread of HIV (Fritz et al., 2002). Alcohol consumption and multiple sexual partners are associated with masculinity, and men that drink heavily often engage in casual sex. Women on the other hand, mentioned in the interviews that they disliked having sex when they have been drinking, often resulting in unwanted sex (Morojele et. al., 2006). Other authors look deeper into the exchange of sex for alcohol consumption in South African drinking venues and show that both men and women understood that accepting alcohol from a man implied consent for sexual favours. Transactional sex and alcohol consumption put both men and women at greater risk for HIV infection trough inconsistent use of condoms and multiple partners (Watt et. al., 2012).

5 Methodology

This paper combines both a qualitative and quantitative method in order to find the effect of the establishment of the Kayelekera Uranium Mine. The qualitative method consists of interviews performed in Malawi during March 2014. The quantitative part consists of analysis of two separate datasets collected in 2004 and 2010.

The qualitative methods are used in order to get a comprehensive view of the HIV/AIDS situation in Karonga and how it has changed since the mine was established. This makes it possible to infuse our knowledge about transmission channels and risky sexual behaviour

from the literature with local issues and infection risks. The quantitative analysis is then used in order to see which of the channels are supported by the data. This enables us to recognize which of the ideas and theories, existing in literature and among the locals, have had a true and significant impact on HIV/AIDS incidence in the area.

5.1 Qualitative Method

The aim of the qualitative method is to get a deeper understanding of the current situation in the District of Karonga and how the establishment of the mine has affected the inhabitants. Specifically, how the mine establishment has impacted HIV/AIDS prevalence through changes in income, alcohol consumption and other potential mechanisms driving HIV incidence.

5.1.1 Choice of Method

Eriksson and Kovalainen (2008) describe that a research question containing “how” aims to answer or explain the causes and consequences of an event or matter in qualitative terms. With a qualitative approach the researcher uses pre-planned questions to start the conversation but let the interviewee take part. Hence, the interview can take many directions depending on the answers given, contrary to a quantitative study. In this study the ambition is to describe how the establishment of the Kayelekera Uranium Mine has affected the individuals in the district of Karonga, specifically the impact on HIV/AIDS and income inequality. Further, Eriksson and Kovalainen (2008) explain that a research question containing “what” aims to explain how individuals experience a process, state or a structural change. In this case the aim is to find out how the key informants interviewed perceive the present and future challenges connected to the Kayelekera Uranium Mine and the spread of HIV/AIDS. Hence, we follow Holstein and Gubrium (2004) and combine “how” and “what” questions, in order to comply with the purpose of the study.

The primary data was collected in the form of interviews, which is used in order to answer the research questions. Guided and semi-structured interviews are used since they, according Eriksson and Kovalainen (2008), are characterized by a prepared outline of topics, issues or themes that allow for the possibility to vary the wording and order of questions during the interview. Further, the authors emphasize that the major advantage of a semi-structured interview is that it can be held in an informal way, with a conversational tone, while the materials are systematic and comprehensive. Semi-structured interviews suit the study well as they allow the respondents to speak freely about how they and people in their surrounding

have been affected by the Kayelekera Uranium Mine. This is also a limitation of the method, since individuals are allowed to have a more open discussion, it can be hard for the interviewer to follow the outline and cover all included topics. However, keeping too close to the pre-planned questions can prevent central topics from being raised by the participant. On the other hand, if the respondents interpret the questions in different ways the empirical materials may be difficult to compare (Eriksson and Kovalainen, 2008).

In order to avoid misunderstandings and bias, the questions we use are neutral and indirect. We started with open ended questions, to encourage more speech combined with some direct questions to stay on the right track. The indirect questions were used since questions regarding HIV/AIDS and sexual behaviour might be sensitive issues to talk about. Finally we used some clarifying questions to be sure that we understand the respondent. Hence, the interviews are built on a combination of various types of questions (Eriksson and Kovalainen, 2008).

5.1.2 Sampling

A non-probability convenience sampling method was used in this study. A limitation of this method is the subjective quality of the assessment and the potential risk of bias within the sample (Kalton, 1983). Regardless of its limitations this method is frequently used in practice, primarily for reasons of convenience as well as cost and time reduction. With respect to our research this method was necessary due to the presence of language barriers and limited access to individuals.

Individuals who were known to live in the area surrounding the mine, primarily from the city of Karonga, were selected. Several respondents, 6 out of 18, were workers recommended by Paladin Africa Ltd, as they were known to be able to communicate in English. This group varied between low and high level workers at the mine site. To counteract the reliance on the mining company further respondents were gathered via local organizations and NGOs e.g. MACRO (Malawi AIDS Counselling and Resource Organization), National Association for People Living with HIV and AIDS (NAPHAM), Banja la Mtsogolo and the US Peace Corps. The organizations were chosen due to their access to and familiarity with the larger population as well as their expert knowledge in areas of public health.

5.1.3 Procedure

The majority of the interviews were held in Karonga, a city located 52 kilometres from the Kayelekera Uranium Mine in northern Malawi. Some interviews were held at the mine site, two in the capital city Lilongwe and two in villages located close to the mine site. Most interviews were held at neutral locations, such as cafés, hotels or, when in the countryside, under a tree apart from the respondents house. No monetary compensation was handed out to the participants; they did however get something to drink and biscuits while answering our questions. Since all participants spoke English, no translator was needed.

The interviews were not recorded as some of the questions might be sensitive to discuss, for instance those regarding HIV/AIDS, sexual behaviour and drinking habits. Instead one of us acted as secretary while the other led the interview. To minimize the loss of important information, we transcribed the interviews on our computers as soon as possible after they were performed.

5.1.4 Data analysis

The aim of the qualitative data analysis is to present the thoughts and sentiments described by the respondents in a systematic and honest way, without including the interview transcripts. This was done through a three step procedure. First the interviews were transcribed, and then the different answers were compared both within each group (the mine-, health administrative workers etc.) and between the groups. Secondly, the key findings were identified and the different opinions are grouped into different categories. Finally the key findings from the interviews were summarized and presented in the structure of a literature review.

5.1.5 Validity

The discussion about internal and external validity is, according to Eriksson and Kovalainen (2008), of great importance when conducting a qualitative analysis. In order to achieve trustworthiness i.e. internal validity, the aim was to interview individuals with insights in the Kayelekera Uranium Mine and how it has affected individuals living in the surrounding area. Further, key-informants from the health sector and NGO`s are included to cover the specific health related questions, more specifically the spread of HIV/AIDS. The interviews consisted of both males and females, even though the males are an overrepresented group when interested in individuals working in the mining sector. In the health sector the gender composition was more equal. We were, however, able to include five females in our interview process.

We have attempted to be systematic and open-minded, when conducting interviews and analyzing the results. There is however always a possibility that we missed out on some important issues, due to our lack of interview experience. Regarding the degree of generalizability i.e. external validity, we believe that when looking at such a specific question in this setting, it is hard to say if the results are generalizable. We do however believe that we can show what can happen when a big foreign direct investment is introduced into a region with high rates of HIV and low levels of development.

5.2 Quantitative Method

Our objective with this paper is to see what has happened to the HIV prevalence rate and income inequality in the Karonga district of Malawi since the establishment of Kayelekera Uranium Mine, which was opened by Paladin Africa Ltd. in April 2009. In order to do so, we have compared data points before and after its establishment. To compensate for the simultaneous development of the rest of the nation, we will compare the change in the Karonga district with the change in a similar district, by using difference in differences-estimation.

5.2.1 Difference in difference-estimation

The basic methodology behind difference in differences-estimation is that an outcome is observed in two different groups for two different time-periods (Imbens and Wooldridge, 2007). The technique is used to measure the effect of an event or a treatment at a given period in time. In our case the treatment is the establishment of the Kayelekera Uranium Mine in the Karonga district. The group is exposed to the treatment between the first and second period; therefore, it is possible to see how the treatment has changed specific variables of interest within the group. Between the two periods, many events may have transpired in the country, or area, which affects the variables of interest, outside of the treatment. Therefore a control group, which shares most traits with the treatment group, albeit not the treatment, is included. If the two groups are sufficiently similar one can assume that the treatment group would have experienced the same development as the control did between the two periods. This is called the common trend assumption and it is vital for difference in differences analysis (Lechner, 2011). The use of two periods takes away the bias of time-invariant factors that might contribute to differences between groups. To find the effect of the treatment, the estimate for the treatment group in the first period is subtracted from the estimate for the second period, this provides the full effect on treatment group. However, in order to compensate for the

changes that are not caused by the treatment, the effect for the control group is also calculated by subtracting the first period-estimate of the group from its second period estimate. The next step is to subtract the control group effect from the treatment group effect, in order to obtain the treatment effect. The treatment effect or the difference in difference-estimator, λ, is equal to:

$$\lambda = (treatment_{t+1} - treatment_t) - (control_{t+1} - control_t)$$

Where time $t + 1$ is the second period and time t is the first. In order to estimate λ econometrically we will use the following expression:

$$y_{st} = \alpha + \beta X_s + \rho T_t + \lambda(X_s * T_t) + \varphi Z_{st} + \varepsilon_{st}$$

Where y_{st} is our variable of interest and X_s is state or district fixed effects that is used to capture the possible differences between the two compared districts, i.e. the control and treatment groups. T_t is a time variable that takes value one for the second period, it captures aggregate factor that would cause a change in y_{st} even without the treatment occurring. The interaction term: $(X_s * T_t)$, is equal to one for the treatment group in the second period, which means that λ is an estimation of the treatment effect. Z_{st} represents a collection of control variables and ε_{st} is the error term.

The advantage of the difference in differences approach is that it controls for group-level- and time-invariant omitted variables. The simplicity of the regression formulation also enables us to easily change the specifications (Angrist and Pischke, 2009), which is useful when performing robustness tests. Possible drawbacks of the technique regard possible endogeneity of the treatment and possible biases that arise from that, as well as the threat of serial correlation of the variables (Bertrand et al., 2004).

Difference in differences-estimators is most commonly performed on repeated cross-sectional data, which means that data is collected on a sample of individuals in the two groups in the first period and then on a different sample in the second period. This puts high pressure on the sampling to be sufficiently random in order to avoid biases. With regards to panel data, difference in differences-estimates is possible, but matching techniques is generally preferred

as causal interpretation of lagged outcomes between the control and treatment groups become problematic (Imbens and Wooldridge, 2007; Lechner, 2011).

In order to find suitable control groups differences in specific characteristic variables between different districts will be analysed. These differences will be checked using t-tests in order to detect significant differences between the compositions among the groups. The proper control group will be chosen on the grounds of similarity to the treatment group, as this implies that the two districts will only differ in terms of being subjected to the treatment. The district must, however, be sufficiently far away from the area of the treatment, so as to not experience any significant spill-over effects.

6 Data and variables

In this section the two different datasets intended for use in the quantitative analysis will be introduced. The source of the data will be presented along with some descriptive statistics.

6.1 DHS-data

In order to analyse the impact of the Kayelekera Uranium Mine on HIV prevalence rates in Karonga a dataset constructed from the Malawian Demographic and Health Surveys (MDHS) of 2004 and 2010 will be used. The 2004 dataset was collected well before the mine was established and corresponds well to how the Karonga district was before construction began. The 2010 dataset was collected slightly earlier than optimal for our thesis purposes, as most data collection occurred in 2009 when the mine had just opened. We do, however, argue that enough migration and behaviour change occurred during the construction phase for the results to be a valid representation of what actually happened in the region. This is based on the peak in the workforce at the mine-site, as well as the peak in HIV incidence for the area (Malawi Government, 2014).

The surveys were primarily performed by the National Statistics Office, and were funded by the government of Malawi, National AIDS Commission and United States Agency for International Development (USAID), among others. Technical assistance was performed through the MEASURE DHS programme, funded by USAID. Three types of questionnaires were used, one for households, one for females and one for males. The questionnaires were intended to provide information about the current population and health issues affecting Malawi. In the 2004 dataset, 458 rural and 64 urban clusters were drawn from the 1998

census sample frame. Out of these, 15,091 households were sampled systematically, in each of these all females aged 15-49 were eligible for the individual interviews, along with a third of the males aged 15-54. In the 2010 dataset, 691 rural and 158 urban clusters were included. These clusters were not dissipated evenly according to population, as an example, the north was slightly oversampled in order to give it a large enough sample size. This is not expected to affect the analysis, as the results are focused on a district and not meant to be generalized for the whole country. For similar reasons, neither weighting nor stratification processes will be used in the analysis. For each district (27 was used, Nkhata Bay and Likoma was combined), at least 950 households was needed in order to provide the appropriate sample size. This gave the dataset a sample of 27,345 households.

What was novel in regards to the 2004 MDHS, and repeated in 2010, was that blood samples were collected from a sub-sample of the overall sample. The blood samples were tested for HIV, thus providing the dataset with a sub-sample including a truthful HIV prevalence. Most HIV prevalence measurements are collected at health centres, which means that the data is collected among people who have got themselves tested. Those types of measures can be biased as these individuals might be more likely both to have knowledge about the epidemic, as well as being more pro-active in their behaviour. Therefore the DHS measure of HIV prevalence is better than most other measurements, provided that the sampling was performed in a correct manner.

In table 1 summary statistics of some of the more important indicators of the Malawian population are presented. The table shows statistics for the whole-sample Malawi including the Karonga district, providing datasets from both 2004 and 2010. There are some notable differences, primarily the fact that the sample size of the 2010 dataset is about double that of 2004. The Karonga observations are more than three times as many in 2010 than in 2004, due to the aforementioned oversampling. Further, the table shows that a larger share of people in Karonga, compared to the rest of the country, are Christian and, generally, have a higher level of education, which shows to be increasing in both datasets. With regards to the wealth index, one can only compare within datasets and not between them, as they are calculated within their respective datasets. However, even with the higher level of education it appears Karonga has become relatively less wealthy than the rest of the country in 2010. The number of children has also increased, as well as access to condoms. The number of observations differs between the variables, since not all are collected on the whole sample (HIV prevalence), there

are missing values (could get a condom) or a prerequisite was needed for the variable (age at first marriage, prerequisite is being married).

Table 1: Summary statistics for the most important variables in the MDHS-dataset

Summary statistics	Malawi		Karonga	
	2004	2010	2004	2010
Current age	27.79	28.19	27.69	28.08
Share rural	0.86	0.87	0.81	0.83
Christian	0.83	0.88	0.96	0.98
Education in years	4.55	5.31	6.37	6.62
Literacy	0.52	0.58	0.58	0.54
Wealth index (DHS)	-0.01	0.01	0.20	-0.13
N.o. children	3.09	3.17	2.79	3.22
Observations	11,420	22,38	202	749
Age at first marriage	17.44	17.42	17.21	17.12
Could get a condom	0.82	0.86	0.54	0.73
Observations	7,867	15,72	159	583
HIV prevalence	0.147	0.122	0.123	0.103
Observations	2,864	7,398	73	252

HIV prevalence appears to have decreased with about two percentage points both in the whole country as well as in Karonga. With regards to Karonga it is hard to tell whether this decrease is due to fewer people contracting HIV or whether people who already have contracted HIV are less resistant to the disease and therefore pass away earlier than in 2004. This is not possible to see here as the table represents the full sample. In the continuing analysis we will only use a specific age cohort in order to estimate incidence, rather than prevalence rates. We have elected 18 to 27-year old women as the cohort to analyse. This is because the life expectancy of a person contracting HIV in Sub-Saharan Africa is about ten years (Mboup et al., 2006). Clinical AIDS seldom occurs earlier than 4 to 5 years after a person gets infected by HIV; it is more likely to occur within 10 to 12 years. After the onset of clinical AIDS the victim's life expectancy is about 1 to 3 years.

In order to construct our own dataset, the variables were matched before merging. The observations for the two districts Likoma and Nkhata Bay were combined in the 2004 dataset, to match the 2010 dataset. Further, the two districts Neno and Mwanza were combined in the 2010 dataset, since the split of the two districts had not yet occurred when the 2004 sample was collected.

Our combined MDHS dataset is a repeated cross-section of data, which, as concluded earlier, makes it suitable for difference in differences-estimation. The individuals' HIV status will be used as the dependent variable. The time variable is the dummy for the 2010 data and the treatment variable will be a dummy variable, taking the value one for Karonga district. In order to account for the treatment effect we will multiply these two dummy variables and create a variable that is equal to unity if the respondent is from Karonga and replied to the 2010 survey. A number of control variables that might affect prevalence rates outside of the treatment effect are also included. Hence our model to estimate is:

$$y_{st} = \alpha + \beta X_s + \rho T_t + \lambda(X_s * T_t) + \varphi Z_{st} + \varepsilon_{st}$$

Where y is blood sample HIV status, β the effect of time, ρ the district fixed effects of Karonga, λ the treatment effect, φ is a vector of the impact of the control variables and ε the error term. This is a fairly straightforward econometric specification, what is needed is to find a proper control group for Karonga.

A suitable control group needs to be one where the parallel trends assumption hold. This infers that the only difference between the control and treatment group will be the treatment effect. Hence it is important to find a control group that is as similar to Karonga as possible, albeit not so close that it experienced spill-over effects of the treatment.

Karonga is situated in the northernmost part of Malawi. The northern region of Malawi is considered to be different than the central and southern regions, as well as in many ways neglected by the government due to its relative remoteness and lack of major cities (such as Lilongwe and Blantyre in the central and southern region respectively). We will therefore focus on the northern region when deciding on control groups.

In order to analyse which district or districts are most fitting as control groups we will look at a number of indicators such as education, religion and fertility. The difference in these indicators between Karonga and our prospective control groups will be tested statistically. Several possible control groups were evaluated, a number of which did not consist of geographically centred groups such as the diaspora of the ethnic group Tumbuka (a large share of people in the Karonga district are of this ethnicity), however we will only present the results of the groups that came closest to being elected.

Table 2 presents summary statistics for all districts in the northern region. It appears that the differences are not that large in most indicators. The Chitipa district is attractive to include in the control group, due to it sharing a long border with Karonga, as well having its main road to Karonga and the rest of the country passing by the Kayelekera Uranium Mine. During the interview process, however, we obtained information that most workers tend to live at the mine-site, in Kayelekera or in Karonga and that little contact exist between the mine and Chitipa. Therefore Chitipa will neither be used as a part of the treatment group, nor as a part of the control group, as there could still be some spill-over effects from the mine. Chitipa also have a large share of its inhabitant coming from other ethnic groups than Karonga, further rendering it inappropriate as control group.

Table 2: Summary statistics for the most important variables in the MDHS-dataset including all the northern districts

Summary statistics	Karonga	Chitipa	Rumphi	Mzimba	Nkhata Bay
Urban	0.80	0.87	0.89	0.76	0.91
Christian	0.98	1.00	1.00	0.99	0.99
Years of education	6.42	6.65	7.46	7.02	6.81
Literacy	0.55	0.625	0.71	0.74	0.70
Number of children	3.11	3.39	3.19	2.90	2.84
Asset index (Wealth)	0.03	-0.01	0.00	0.26	0.09
Age at first marriage	16.78	17.37	17.93	17.25	17.49
Access to condoms	0.80	0.69	0.70	0.65	0.94
HIV prevalence	0.1078	0.0429	0.0749	0.0794	0.1293

The significant differences between Karonga and the four possible control districts, stemming from t-testing, are presented in table 3. The differences are split up into significance levels of the differences. As mentioned in the paragraph above, Chitipa is precluded, as it is ethnically different and might have experienced spill-over effects from the mine. Rumphi is also excluded from our analysis; both due to the low number of observations, in 2004, as well as the amount of significant differences (e.g. Rumphi had no respondents from urban areas).

Table 3: The most important results from t-tests performed on the MDHS-dataset

Differences	1%	5%	Observations
Chitipa	Wealth	Urban	40
Rumphi	Urban, Wealth	Literacy	34
Mzimba	Urban	Education, Literacy	241
Nkhata Bay	-	-	53
Mzimba/Nkhata Bay	-	-	294

People in Mzimba are significantly more educated and a smaller share of them live in urban areas. When comparing Nkhata Bay and Karonga no significant differences are found, there are however, few observations. Therefore, a group consisting of Mzimba and Nkhata Bay is constructed. These two districts share a long border and together stretch from the border to Zambia to the shore of Lake Malawi in the southern part of the northern region. This combined group have a large number of observations and is not significantly different from Karonga in any of our indicators. Both districts also have the benefit of not sharing a border with Karonga, thus making it unlikely that they received any spill-over effects from the establishment of the Kayelekera Uranium Mine. Therefore, a combination of Mzimba and Nkhata Bay is used as the control group in our analysis.

6.2 LSMS-data

In order to analyse income inequality we need data on incomes, which is a significant problem in Sub-Saharan Africa as the monitoring of individual revenues is not sufficient to be used in data analysis. Many people do not work in the formal sector, and if they do it is not always the case that they receive their payments in a manner that can be easily measured. A large share of the population is also engaged in subsistence farming, which often does not translate into hard numbers on a bank account. In this paper a measure of wealth will be used in order to see what has happened since the Kayelekera Uranium Mine opened.

In order to analyse how the Kayelekera Uranium Mine has affected income inequality, we will use data from the Living Standards Measurement Studies (LSMS) of the World Bank. The Integrated Household Surveys, performed by the LSMS project in 2004 and 2010, are rich datasets that contain information about household assets, welfare, health and education. The surveys were designed by the National Statistics Office (NSO) of Malawi with technical assistance from the World Bank and International Food Policy Research Institute (IFPRI). The purpose of the surveys is to provide a complete and integrated dataset to better understand which groups in society are most vulnerable to and affected by poverty. The dataset is to be used as indicators, to assist policymakers in their attempts to reach the Millennium Development Goals and, specifically in Malawi, the goals listed in the Malawi Growth and Development Strategy.

In our analysis we will use the second Integrated Household Survey (IHS2), collected in 2004, and the third IHS (IHS3) from 2010. IHS2 was assembled before the mine was

established and corresponds to how the income distribution in the district of Karonga was prior to the mine establishment. IHS3 was collected one year after the mine officially opened. Our purpose is to use this data to investigate the differences before and after the mine establishment. The data collected in 2010 might not be optimal when investigating the effects the mine has had on the district, as the mine had just opened. However, we argue that a large share of the in-migration, both of professionals and people looking for work, and economic activity occurred during the construction of the mine, between 2007 and 2009. Therefore, much of the economic impact of the area is included in our dataset. A follow-up study was performed in 2013, of which the data is unfortunately not available until September 2014 and can therefore not be used in this paper.

The IHSs are far-reaching multi-topic surveys done over a period of twelve months. In both IHS2 and IHS3 a two-stage stratified sampling procedure is used. The primary sampling units were enumeration areas (EA), and in the second stage a random selection of households in each EA was elected to be surveyed. The 2004 survey consists of 11,280 households and the 2010 survey contains of 12,271 households. The two datasets are independent of each other, but contains comparable variables.

To ensure comparability across districts and time, only variables that appear in both questionnaires are included in the analysis. Hence, the two datasets have been matched, extracting the variables of interest that existed in both surveys. The variables are cleaned and then transformed into binary variables. The districts in the third dataset have been changed to match the second one and the two datasets have then been merged. This new dataset consist of a number of different consumption goods, such as charcoal, matches, mosquito nets, beds, tables, chairs etc. as well as time and district variables.

To estimate wealth we will construct an asset index, which will account for the possession of a number of specific goods, where a higher number of goods implies higher wealth and standard of living. The perk of an asset index is its simplicity, as it is a clear indicator of material well-being (Young, 2012). Sahn and Stifel (2000) point out the added benefit of not needing to use price deflators and currency convertors, which might be biased, due to inflation or other events regarding monetary policy and price levels. There are a number of pitfalls when using asset indices. Asset preferences and relative prices might change over the years and can change the subjective importance of specific assets (Harttgen et al. 2013). It

might also be problematic using an index based on stock variables when attempting to estimate a flow variable, such as income. Specifically this might be an issue when the age and model of the goods is of importance.

Specific goods might influence the variation in wealth to a different extent. Therefore, the use of a simple accumulative index of binary variables is not able to pick up the subtleties of a household's wealth. For that reason multiple correspondence analysis will be used when creating the asset index. Multiple correspondence analysis detects underlying structures in nominal categorical data and will mirror wealth if applied to our dataset of assets (for further reference, see Appendix B).

7 Results

In this section the results from the qualitative and quantitative analysis will be introduced. First are the qualitative results followed by the quantitative results, the latter consisting of the analysis of changes in HIV prevalence and risky sexual behaviour using the DHS-dataset, followed by the analysis of wealth inequality using the LSMS-dataset.

7.1 Qualitative findings

The major findings from the interviews will be presented below. The first part will focus on how the Kayelekera Uranium Mine has affected the individuals living in Karonga and the surrounding villages when it comes to income levels, inflation and income inequality. The second part will focus on how the spread of HIV/AIDS has been affected since the mine was established.

The general impression from the interviewees is that the city of Karonga and the surrounding area have changed significantly since the mine was established. The respondents speak about how the construction of houses has developed, from mud and wood, to houses made out of bricks. There are now more individuals who can afford to buy, for example, cars and the road infrastructure and communications have improved. There are now several banks operating in the area and some respondents also highlight the increase in small businesses in the city of Karonga. There are businessmen from Tanzania, China and India who have established businesses selling such things as clothes, shoes and hardware. The number of bars, commonly referred to as boozing dens, has increased and there are now more places to find different kinds of amusement.

The direct effect of the opening of the Kayelekera Uranium Mine was that the company offered relatively high wages to attract skilled labour. This lead to positive income shocks affecting both the domestic and migrant workforce. These income shocks were sometimes hard to handle. A male mine-working respondent said "*When people get their salary, they just think about enjoyment. And once you work for the mine, people think that you are the 'big boss' and you are supposed to spend a lot of money. "* The workers usually have a short time horizon and have a tendency to spend their money on different types of enjoyment, such as alcohol and sex. There is also a multiplier effect present, due to interdependence. Individuals working are expected to support others, as their extended family. This is both the social norm and not all too surprising as most individuals living in the Karonga district are self-sufficient farmers that are not integrated in the formal economy.

There are also indirect income effects, which arise from the demand of different types of goods and services from the mining company. Companies located in Lilongwe and local companies in Karonga have been able to develop, as the demand for items such as bananas, maize and other products/services have increased.

The relatively high salaries have also caused the price of housing, food and other commodities to rise. Several respondents shared the same experience: individuals working for the mine have the ability to overbid others. A male, mine-working respondent said "*Before the mine came it would cost you 15000 MWK to rent a good house, but now someone can come and ask how much the current tenant is paying and offer 25000 MWK for the same house."* This implies that the current tenant needs to leave. Some also state that vendors set higher prices if they know that a person is working for the mine i.e. they practice price discrimination. The general understanding is that all the prices have increased, both when it comes to staple goods, such as maize and fish, as well as housing. This price increase is experienced both in the city of Karonga and the villages surrounding the mine. Since this problem arose, the mining company has re-organized their purchasing to a few local suppliers and do not contribute to the price increases in the same manner as before. Positive aspects that two respondents mentioned is that the availability of goods has increased since the mine came. In addition some of the mine workers invest in smaller businesses and hire employees. Hence, many people benefit from the high salaries earned.

The respondents also mentioned that income inequality has increased in the region. The inflation has made some people with the same income levels, as prior to the mine, relatively poorer, while people benefitting from the mine have become richer. A new class has emerged in the area, since more people are earning higher salaries. This is however, not only caused by the Kayelekera Uranium Mine, there are also other local businesses that have aided the development. A positive effect of the mine, as some respondents mentioned, is the fact that some mineworkers open up small businesses or employ people working as housekeepers and cleaners. A male respondent expressed that the already rich became richer when the mine came, as it was the individuals with already existing businesses that benefitted from contracts with the mine in the first phases of development. The wealth has therefore not been equally spread in the community.

All the respondents expressed that, in general, people know about HIV/AIDS and how it is transmitted, however, there is still stigma associated with the decease. Individuals are not open with their HIV-status; instead people usually keep it to themselves. Many respondents also mentioned that the individuals who do go and test themselves are usually the ones who believe that they are not infected. In addition, few individuals attend regular check-ups. The general belief is that one test is enough, while the health sector recommends the individuals to get tested once every three months.

Two of the respondents working in the health sector express that several experts predicted that HIV prevalence rates would increase significantly in the area when the mine was established. During the construction phase in 2007, there were a lot of migrant and local workers earning a good salary. The migrant workers were away from their families and many lived in the Kayerekera village. This attracted a lot of sex workers, both from Malawi and neighbouring countries such as Tanzania. There even existed sex for credit during this phase, according to a male working at the mine. But after the construction phase, as the number of workers decreased, the number of sex workers has also declined, specifically in the Kayerekera village. According to other respondents it is clear that prostitution in Karonga still persists. Gender inequalities are a contributing factor to the large number of women going into prostitution. A female respondent, working in the health sector said: *"After working at the mine for a month, the males come into town (Karonga) to spend their money on prostitutes instead of giving money to their wives and children. So, what are the women supposed to do?"* The gender inequalities have a positive impact on the number of sex workers. Many

respondents say that poor women have no other chance to support themselves and their children than engage in transactional sex. There are also some groups of individuals that are at higher risk than others to contract HIV/AIDS. These groups include fishermen, market vendors and sex workers. The fishermen leave their partners behind when away fishing and, while away, both the fisherman and in some cases their wives might take other partners. This increases the spread of HIV further. When the fishermen return, they sell the fish and often receive a big lump sum cash payment. This money is often spent on alcohol consumption and transactional sex. Two respondents mentioned that females sometimes use their bodies as payment in order to buy, for example, fish.

The higher salaries and differences in wealth have also contributed to increased alcohol consumption in the area. According to 11 out of 18 respondents, alcohol is one of the major drivers behind the spread of HIV and AIDS in the region. When people are drunk they seem to disregard their knowledge about HIV and engage in unprotected sex. One respondent said "*People that are drunk even have condoms in their pockets.*" Another respondent said "*The people that get money often invite their friends and buy drinks to everyone. And when people are drunk, they are more likely to engage in risky sexual behaviour and take multiple partners.*" Before the mine came, it was common to brew your own beer, a practice that now is illegal. It is, however, still very easy to get your hands on cheap spirits, specifically in the form of sachets, one decilitre spirits that cost around 30-50 MWK (around €0.06-0.09). This abuse have reached young boys as well, not older than 12-13 years. The increased alcohol consumption is driven by males, but many females engage as well.

There are also other factors driving the HIV/AIDS epidemic. In northern Malawi, polygamy is common. Most respondents say that it is common to have more than one wife, having up to four wives is allowed. One respondent said "*You do not have to be rich to have many wives. The males want to show off, show that they are a popular man, having many wives and children.*" It is also common to have girlfriends on the side. These trends are, according to our respondents, more persistent in the rural areas. Intergenerational sex is also a problem, older men together with younger girls. Older men pay the girls for sex, but once they get pregnant the males leave them and they have to drop out of school, without having anyone that can support them. The cultural phenomena "Kupimbira" is another practice, meaning that a man is betrothed to a young girl and is able to marry her at a time of his choosing. There is the "Kuhara" where a husband's brother "inherits" the deceased brother's wife as well as

practices where a man taking a wife can also take her younger sister as an additional wife. Many respondents say that these practices are declining as people leave their traditions and get more knowledge about how these practices can affect the spread of HIV/AIDS.

Some respondents say that other types of superstition still exist; most do however believe that it is more common in the rural areas than in the urban areas. There are still some people that believe that you can get rid of HIV by sleeping with a virgin or, alternatively, sleeping with a lot of different partners. The respondents do believe that the spread of information regarding HIV/AIDS in recent years has decreased this superstition significantly. The key informants from the health sector also mention that the demand for condoms has increased, indicating that individuals are actually listening to the messages about safe sex, spread by different groups in the community.

7.2 Quantitative findings

In this section the results from the quantitative analysis will be presented. It will start with the results from the DHS-data analysis of risky sexual behaviour and continue with the LSMS-dataset and the analysis of wealth inequality.

7.2.1 DHS regression results

Table 4 presents the baseline difference in differences-regression results; a logistic or logit regression method is used since our dependent variable is binary. The first column shows our baseline regression with the Karonga district as the treatment group and Mzimba together with Nkhata Bay as the control group. The regression is performed on data collected on women aged 18 to 27. The results show no significant treatment effect, which would have indicated that HIV-incidence changed more in Karonga than in Mzimba and Nkhata Bay. The results do however; indicate that women in urban areas are more likely to get infected than women living in rural areas. There is also a significant age effect which means that the probability of contracting HIV is larger the older a person in this specific age cohort is.

The issue that occurs when using this data and type of specification is the problem of the sample size of Karonga, specifically in the 2004 part of the data. When the data is scaled down into the 18 to 27-year old cohort, a large share of the data is lost and this decreases the amount of HIV-positive people in the sample to four. The first regression relies significantly on these four HIV-positive individuals and it biases the results to a large extent. Therefore in

Table 4: Baseline regression on HIV using difference in differences-estimation

	(1)	(2)
Dependent variable	HIV	HIV
Treatment	0.614	0.455
	(0.690)	(0.420)
Time	-0.568*	0.205***
	(0.343)	(0.0762)
Treatment*Time	-0.169	-0.523***
	(0.371)	(0.102)
Education	0.117	0.106**
	(0.0762)	(0.0500)
Urban	1.182***	1.146***
	(0.419)	(0.164)
Age	0.0954**	
	(0.0373)	
Age 24-28		0.478**
		(0.217)
Age 29-33		0.690
		(0.480)
Age 34-38		1.434***
		(0.422)
Age 39-43		0.983***
		(0.128)
Age>43		1.628***
		(0.480)
Constant	-5.592**	-3.867***
	(2.255)	(0.941)
Observations	441	910

Clustered standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

column (2) the full sample data of Karonga is included, i.e. all ages. This is however, a trade-off; when all age groups are included we get more observations, but, at the same time, a higher amount of individuals have been infected with the HIV virus before the mine establishment. This means that only prevalence and not incidence can be analysed, which complicates interpretation as the change might have occurred because more people contract HIV or people with HIV live longer and thus represent a larger share of the population. The results show a significant treatment effect which tells us that HIV prevalence decreased in Karonga as a result of the treatment. In the table the logistic regression outcomes are displayed, not the margins, therefore it is not possible to comment on the magnitude of the effects and only on the sign and significance. Instead of using age as a regressor the sample is

split into six age cohorts, first there is people younger than 24 and then cohorts for each five-year intervals until the last cohort which consists of people over 43. HIV prevalence is higher for the older cohort as well as for the 24-28 age cohort. It is also higher for people in urban areas and people with more years of education.

Once again, the dataset is plagued with overreliance on a too limited sample of observations. Table 5 presents the total number of observations of HIV status in our control and treatment groups. In the 2004 dataset there were only 8 recorded cases of females being infected with HIV and in order to make proper statistical inference this number ought to be at least 30. In our case, there are not 30 observations for neither of the districts nor for any of the time periods, except Mzimba/Nkhata Bay in 2010. Therefore we cannot base any conclusions about the effects of the Kayelekera Uranium Mine on HIV prevalence rates in Karonga, based on evidence from this dataset.

Table 5: HIV-status, number of observations divided by treatment, control group and year

	Karonga		Mzimba/Nkhata Bay	
HIV status	2004	2010	2004	2010
Negative	48	172	265	475
Positive	8	23	29	52

We will instead look at changes in indicators of risky sexual behaviour in Karonga. This will be done by, first, looking at the full Malawi HIV dataset in order to analyse what indicators of risky sexual behaviour have a significant effect on HIV prevalence rates. Secondly the full MDHS-dataset will be used. This dataset has the perk of having about three times more observations than the HIV dataset, however it lacks the most important indicator, HIV, as that analysis was only performed on a subset of the sample population. In this dataset difference in differences-estimation will be applied with the significant indicators of risky sexual behaviour as dependent variables in order to see if this behaviour has changed in Karonga and can help explain the 2009 peak in HIV incidence, as evidenced in the data provided by the Ministry of Health.

Table 6 columns (1), (2), and (3) display the significant results for a regression of a number of indicators of risky sexual behaviour on HIV status, for the whole country, including all available age groups. In column (1) a logistic regression is performed using indicators of risky sexual behaviour as regressors of the dependent variable HIV. There are significant positive

Table 6: Statistically significant regression results including a number of indicators of risky sexual behaviour

	(1)	(2)	(3)
Dependent variable	HIV	HIV	HIV
Urban	0.859***		0.800***
	(0.145)		(0.133)
Age-gap	0.0192***	0.0200***	0.0188***
	(0.00716)	(0.00714)	(0.00720)
Virgin-bride	-0.529***	-0.558***	-0.532***
	(0.106)	(0.108)	(0.106)
Alcohol	0.250*	0.252*	0.251*
	(0.129)	(0.135)	(0.132)
Teenage pregnancy	0.0754***	0.0720***	0.0756***
	(0.0188)	(0.0191)	(0.0188)
Rich		0.452***	0.218
		(0.163)	(0.160)
Constant	-2.451***	-2.412***	-2.469***
	(0.211)	(0.218)	(0.230)
Observations	6,695	6,695	6,695

Clustered standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

impacts on HIV prevalence if the respondent lives in urban areas, their partner gets drunk often, the age gap between the respondent and its partner, as well as if the respondent had a child while being a teenager (this effect is linear and increases the younger the mother was). There is also a significant negative effect on HIV status for women being virgins until they get married. Column (2) include wealth status and the result indicate that rich people tend to have a slightly higher HIV prevalence as well as respondents with less dominant partners (where the partner makes all decisions about money in the household). Column (3) includes the dummy variable for urban areas again and the wealth effect of column (2) vanishes. This leaves us with five tangible and significant indicators of the effect of sexual behaviour on HIV status. (Full results are available in Appendix A – Appendix table 2).

HIV prevalence tends to be higher in urban areas due to higher population density as well as higher levels of migration than in rural areas. High consumption of alcohol also has a positive impact on HIV prevalence, as inhibitions tend to decrease when a person gets drunk. Unprotected sex and coital frequency are also likely to increase with alcohol consumption. A large age gap within couples (male older than female) tends to increase HIV-prevalence as it induces females to have their sexual debut earlier, it gives more power in the relationship to

the male, it often results in less education for the females as they are sometimes pulled out of school if they get married early, these males are also likelier to be polygamists and the possibility that the male's first wife died of AIDS means that the male is likely also infected with HIV. Giving birth while in one's teens indicates an early sexual debut as well as higher likelihood of being engaged in a long-term relationship. When a girl gets pregnant she is usually pulled out of school and is forced to depend on others to a large extent. These girls usually marry early and in many cases, especially the younger the girl is, she is taken as a second wife. Less schooling, early sexual debuts and polygamy all indicate a higher HIV prevalence. Females being virgins until marriage has a negative effect on HIV prevalence simply as these women are not having intercourse before marriage, thus taking away the most important transmission channel: sex. This might also display a specific personality trait, which makes these women less likely to give in to temptation, such as premarital sex and infidelity.

Table 7 columns (1), (2), (3) and (4) shows the outcomes of logistic difference in differences-regressions on four of the significant indicators of sexual behaviour. This is done in order to see whether risky sexual behaviour has increased in Karonga. The results, however, generally do not display any stark changes in risky sexual behaviour in Karonga. Two of our indicators have a significant treatment effects. The likelihood of a female being a virgin until marriage in Karonga is spectacularly higher since the mine opened. The treatment effect indicates that women in Karonga are about 88% more likely to be a virgin until marriage in 2010 than in 2004 (see table 8). In this regard however, most of the explanation stems from the extremely low levels of the 2004 data, where 1.49% of the women in Karonga remained virgins until their first union, comparable to a rate of 35.25% in 2010, which is closer to the national average of about 25%. This result implies that HIV prevalence in Karonga ought to have decreased, as evidenced in table 6. This result is in line with Wilson (2010) who found that the opening of a mine will decrease most indicators of risky sexual behaviour, however it only applies to one of the estimators. There is no significant effect of the age-gap when the test is performed for women who have married since construction of the mine started (married later than 2006), i.e. those who might have been affected by the establishment. This modification is used both for virgin until first union and age-gap, so as to not bias the results with those who were married before the effects of the mine had appeared. There is no change in the probability of a respondent's partner getting drunk since the mine was established. There is,

Table 7: Difference in differences-estimates on four indicators of risky sexual behaviour

Dependent variable	(1) Age-gap	(2) Virgin-bride	(3) Alcohol	(4) Teenage pregnancy
Treatment	1.125***	-3.932***	0.329***	0.0236
	(0.0428)	(0.251)	(0.0815)	(0.137)
Time	0.852	-0.389***	-1.167***	-0.380**
	(1.042)	(0.138)	(0.147)	(0.0830)
Treatment*Time	-2.025	5.123***	-0.0912	0.406*
	(0.894)	(0.163)	(0.140)	(0.123)
Constant	6.613***	-1.907***	-2.760***	2.236***
	(0.567)	(0.538)	(0.362)	(0.158)
Observations	1,174	1,599	3,720	2,033
R-squared	0.011			0.110

Clustered standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

however, a significant change, albeit at 10% confidence level in teenage fertility as a result of the treatment. In this case only those who were teenagers in 2007 when construction started are included, so as to not include anyone who's teenage fertility was unaffected by the mine. The treatment effect is positive which indicates that more women are getting pregnant in their teens than before. As noted in table 6, the increase in teenage fertility ought to increase the incidence of HIV in Karonga. Table 8 show that the marginal effect of teenage fertility is larger than one, which is calculated on the derivative of the logistic regression. Noticeable about table 7 is that both the age-gap and alcohol consumption have negative signs, albeit not significant results. (Full results are available in Appendix A – Appendix table 3).

Table 8: Marginal effects of difference in differences-estimates on Virgin-bride and Teenage Fertility

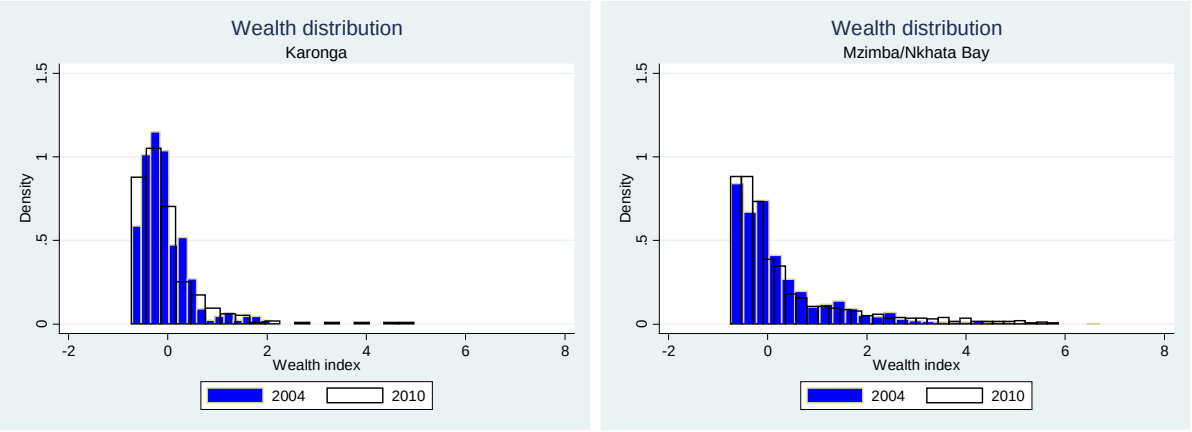
Marginal effects	Virgin-bride	Teenage Fertility
Mzimba/Nkhata Bay	0.2429	1.3693
Karonga	0.8785	1.7756
Marginal effect of treatment	0.6356	0.4062

With regards to the impact of the urban variable, a t-test was performed in order to see whether the share of people living in urban areas had changed in Karonga. The results were not significant, which implies that the ratio is stable and the increase in HIV was not caused by a higher level of urbanization. (See Appendix A – Appendix Table 4).

7.2.2 Income inequality

With regards to the change in equality we developed an asset index using multiple correspondence analysis. Figure 2 displays the distribution of the asset index for the treatment and control group divided by year. The blue bars represent the distribution of income in 2004 and the white bars the distribution in 2010. It appears that the share of people in the lowest percentile is larger in 2010 than in 2004 implying that the share of relatively poor people is higher. The difference is not as clear when wealth is higher, there is a larger share of people with a wealth index value of about 0.5 in 2004 and similarly a larger share in 2010 with wealth at about 0.2. From this our expectation is that income inequality in Karonga will have increased due to more people being in the poor segment.

Figure 2: Distribution of the asset index for the treatment and control group



The control group display a similar distribution, again it appears as a larger fraction of the population is in the poorest share in 2010, however, the difference is not as large as in the treatment group. The number appears to have increased significantly in the second poorest group, this combined with the first result implies that the share of poor people have increased in Mzimba and Nkhata Bay. In this graph it also appears as if a larger segment of people is rich in 2010 as when the index is larger than about 2.5 almost all white bars are higher than their blue counterparts. From this visual evidence we expect income inequality to have increased in the control group as well.

Table 9 presents the mean values of wealth in the tenth and ninetieth percentile, per district and time. The larger the difference in wealth between these percentiles, the higher the income inequality will be within the district. The table shows that the means in the tenth percentile is

Table 9: Difference in wealth between the 90th and 10th percentile

	Karonga		Mzimba and Nkhata Bay	
Time	2004	2010	2004	2010
Lowest decile	-0.6684	-0.6679	-0.6852	-0.6660
	(0.0077)	(0.0092)	(0.0044)	(0.0030)
Highest decile	1.4530	2.0690	2.1969	2.5767
	(0.0991)	(0.2496)	(0.0843)	(0.0849)
Difference	2.1215	2.7369	2.8821	3.2427
	(0.0994)	(0.3650)	(0.1688)	(0.1575)
Welch t-statistic of difference	0.6154		0.3605	
p-value of difference	<0.0001		<0.0001	
Observations	31	22	205	266

similar in both districts, indicating that the poorest ten percent of people in both districts have similar wealth and living standards. When looking at the ninetieth percentile, however, the value for Mzimba and Nkhata Bay is higher in both time-periods, implying that the rich people are richer in Mzimba and Nkhata Bay and that income inequality is higher than in Karonga. How has the inequality changed then?

In order to look at the difference a Welch's t-test is used, which is an adaptation of the Student's t-test. The test allows for hypothesis testing between two populations with different variances. The results are summarized in table 9 (calculations available on request). From here it is possible to deduct that income inequality has increased in both districts between 2004 and 2010. It is also noticeable that the change in Karonga is larger than that in Mzimba and Nkhata Bay, suggesting that the rate of increasing inequality has been higher. Mzimba and Nkhata Bay is still more unequal, albeit Karonga is catching up. We are unfortunately neither able to test the difference between the districts nor perform a difference in differences. However, the large difference (70.7% higher increase in Karonga) implies that the establishment of Kayelekera Uranium Mine has resulted in larger differences in wealth among the people in Karonga.

Next we turn to the use of Gini-coefficients in order to see how wealth inequality has changed. The asset index consists of a distribution of values with a mean of zero. This means that there are observations with negative values of wealth. In order to calculate Gini-coefficients properly the asset index must be transformed into only positive values. Therefore,

Table 8: Difference in Gini-coefficients between treatment and control group

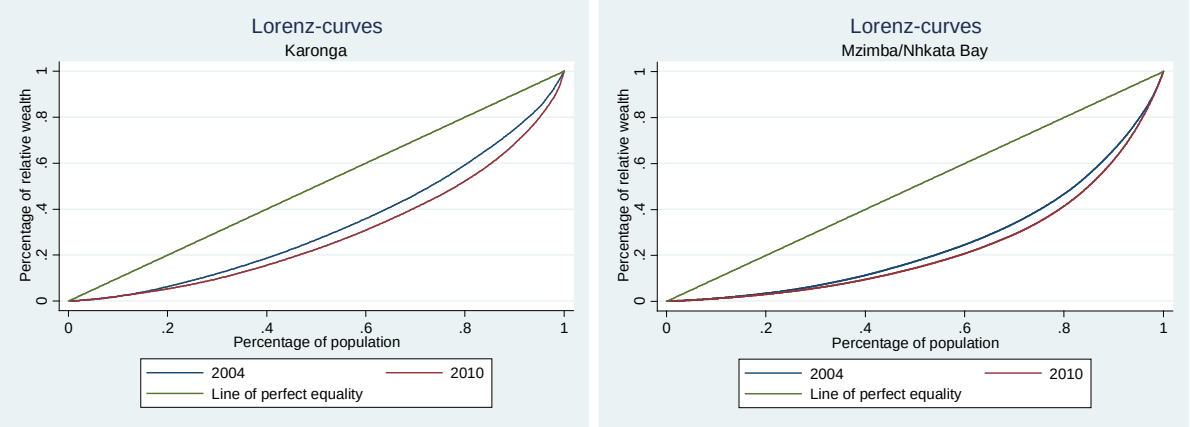
	Karonga		Mzimba/Nkhata Bay	
Time	2004	2010	2004	2010
Gini coefficients	0.3427	0.4149	0,4878	0,5359
	(0.0169)	(0.0184)	(0.0088)	(0.0072)
Difference	0.0722		0.0481	
p-value of difference	<0.0001		<0.0001	
Observations	240	380	960	1150

we follow le Roux Booysen et al. (2005) and add the largest observed negative value to each of the asset index observations. The lowest observed value becomes about zero when adding 0.7654292 to the asset index. This transformation does affect the resulting poverty and inequality measures, meaning that the values have no meaning on their own. Nevertheless, even if the values has changed per se, the distribution of the values are the same as before the transformation and the values still has a meaning in a relative sense.

The transformation makes it possible to compare the estimated asset index across time. The t-test are presented above, in table 10, and shows that there is a statistical significant difference in income inequality in the both the treatment and control group. The income inequality is higher in 2010 compared to 2004, in both groups and the income inequality has increased more in Karonga than in Mzimba and Nkhata Bay, this difference can unfortunately not be tested, however. As mentioned above, these results can be biased because of the transformation process.

The transformed asset index is also used to create Lorenz-curves in the treatment and control districts. These are presented in Figure 3. The graphs show that Karonga has a lower inequality than Mzimba and Nhkata Bay to start with. Both the control and treatment groups have experienced higher income inequality in 2010, compared to 2004, and it appears as the change in Karonga is larger than the change in Mzimba and Nhkata Bay.

Figure 3: Lorenz curves for treatment and control group



8 Robustness

In this section the robustness of the quantitative results will be evaluated. The analysis of risky sexual behaviour, using DHS data will be tested using a dependent variable other than HIV blood sample status: STIs which are related to HIV/AIDS in several ways. The robustness of the analysis of income inequality will be checked by using another asset index, calculated form the DHS-dataset, rather than the LSMS-dataset.

8.1 Sexually transmitted infections

Our first robustness check will be to perform the HIV blood sample status regressions using STIs (Sexually Transmitted Infections) as the dependent variable. HIV in Sub-Saharan Africa is mainly spread through sexual intercourse and can therefore be considered an STI. STIs are a good indicator of risky sexual behaviour as they are spread mostly through unprotected sex. If, for example, a person have many partners, sex frequently or drinks heavily the risk of contracting an STI increases, in a similar fashion as for HIV. In the Malawian Demographic and Health Surveys of 2004 and 2010, variables measuring self-reported STIs and STI symptoms are included. The first variable shows whether the respondent have been diagnosed with a STI in the last 12 months before the interview. This means that only people who have visited a health clinic in the last year are eligible for a positive value for this variable, limiting the sample. The second variable measures whether the respondent has experienced any of the two most common symptoms of STIs; a bad smelling, abnormal discharge from the vagina or penis or a genital sore or ulcer, in the last year. If any of these symptoms have been experienced it is an indicator of a STI, albeit not diagnosed or treated yet. The second variable, symptom, will be used as a proxy for risky sexual behaviour and assessed whether it has changed in Karonga or not.

Table 9: Difference in differences-estimates on Sexually Transmitted Infections

	(1)	(2)
Dependent variable	STI	STI
Treatment	0.0592	-0.503
	(0.472)	(0.511)
Time	0.0498	0.589
	(0.835)	(0.759)
Treatment*Time	-0.163	-0.313
	(0.808)	(0.748)
Education	-0.0347	-0.0116
	(0.0420)	(0.0225)
Urban	-0.584	-0.108***
	(0.412)	(0.0118)
Age	-0.0113	
	(0.0547)	
Age 24-28		0.0848
		(0.418)
Age 29-33		-0.0855
		(0.601)
Age 34-38		0.261
		(0.497)
Age 39-43		0.537
		(0.469)
Age>43		-0.172
		(0.588)
Constant	-2.770***	-3.524***
	(0.678)	(0.517)
Observations	1,376	3,047

Clustered standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 11, display a logistic difference in differences-regression similar to the one in table 4, with the only difference that now STI symptoms are the dependent variable. Column (1) displays a statistically significant negative interaction term, implying that the number of people displaying STI symptoms decreased in Karonga compared to the control districts, Mzimba and Nkhata Bay. However, these results are unfortunately not reliable as the same issues as when HIV was used as the dependent variable are present. There are simply too few observations of STI symptoms in Karonga in 2004 to draw any conclusions. The results are biased and interpretation would be misleading.

Table 10: Regression on Sexually Transmitted Infections including a number of indicators of risky sexual behaviour

Dependent variable	(1) STI	(2) STI	(3) STI
Education	-0.0197*	-0.0222*	-0.0230**
	(0.0111)	(0.0115)	(0.0113)
Age-gap	-0.0103*	-0.0103**	-0.0104**
	(0.00526)	(0.00525)	(0.00526)
Virgin-bride	-0.337***	-0.338***	-0.337***
	(0.0575)	(0.0569)	(0.0575)
Alcohol	0.222*	0.222*	0.221*
	(0.114)	(0.118)	(0.117)
Polygamy	0.118**	0.115**	0.117**
	(0.0493)	(0.0490)	(0.0493)
Teenage pregnancy	0.0316**	0.0321**	0.0321**
	(0.0133)	(0.0133)	(0.0133)
Working	0.183**	0.182**	0.183**
	(0.0807)	(0.0802)	(0.0804)
Dominance	0.244***	0.245***	0.247***
	(0.0749)	(0.0750)	(0.0748)
Infidelity	0.659***	0.653***	0.654***
	(0.174)	(0.176)	(0.176)
Constant	-2.211***	-2.189***	-2.191***
	(0.114)	(0.122)	(0.123)
Observations	22,864	22,864	22,864

Clustered standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Therefore we take a similar approach as that of the HIV regressions. Columns (1), (2) and (3) in table 12 report the significant results from a logistic regression on the whole country, again using STI symptoms as the dependent variable and including various regressors. This is done in order to identify any variables that might significantly impact the risk of showing STI symptoms i.e. engage in risky sexual behaviour. Similar to the regression on HIV, there is a significant positive impact on facing STI symptoms if the respondent has experienced a teenage pregnancy and a negative impact if the respondent were a virgin until marriage. The age gap between partners has a significant negative impact on STI symptoms, which contradicts the results obtained from the HIV regressions. This result implies that as the age-gap increases the respondent is less likely to contract a STI. Thus, the result is counterintuitive and cast doubt about whether our earlier results with regards to age-gap are robust or not.

Similarly to the HIV regressions here alcohol has a significant positive impact on the likelihood of a respondent contracting a STI. This means that a person is more likely to be infected by a STI if the respondent's partner often gets drunk. In addition, the results show that there is less risk of showing STI symptoms if the respondent is educated, as well as higher risk if the respondent is engaged in a polygamous relationship, has a dominant partner or if the respondent is working. The results are maintained independent of the different regressors included and give us eight significant indicators that are related to risky sexual behaviour and the probability of experiencing STI symptoms. (Full results are available in Appendix A – Appendix table 5).

In table 7 we displayed the change in four of the significant indicators (age-gap, virgin-bride, alcohol and teenage fertility) for the Karonga district and in table 13 we perform the same type of difference in differences-estimation for the remaining four indicators (polygamy, working, dominance and infidelity). The interaction term is positive and significant for only one of these indicators: dominance. The share of people who have dominant partners have increased, these results are however, only significant at a 10% significance level. The dominance variable is a dummy that takes value one if the husband makes all decisions regarding the expenses of the household. An increase in the number of women with dominant husbands ought to increase the spread of HIV, as evidenced in table 12. If the husband is very dominant, he is more likely to suppress his wife and more likely to have more sexual partners. Table 14 displays the marginal effects of the dominance estimate, the likelihood of a woman

Table 11: Difference in difference-estimates on four indicators of risky sexual behaviour

	(1)	(2)	(3)	(4)
Dependent variable	Education	Polygamy	Working	Dominance
Treatment	-0.298	0.293***	-1.069**	-0.0261**
	(0.112)	(0.00359)	(0.513)	(0.00268)
Time	0.433**	-0.332*	-0.524	0.0282
	(0.0918)	(0.175)	(0.860)	(0.0369)
Treatment*Time	-0.0987	0.138	0.512	0.131*
	(0.0914)	(0.167)	(0.859)	(0.0365)
Constant	8.682***	-2.124***	-0.819*	-0.0213
	(0.187)	(0.0822)	(0.447)	(0.0267)
Observations	3,720	2,571	3,720	3,720
R-squared	0.180			0.042

Clustered standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 12: Marginal effect of difference in differences-estimates for dominance

Marginal effects	Dominance
Mzimba/Nkhata Bay	0.0389
Karonga	0.3985
Marginal effect of treatment	0.3596

having a dominant partner is 35.96% higher in Karonga than in Mzimba and Nkhata Bay, since the mine establishment. (Full results are available in Appendix A – Appendix table 6).

8.2 Income inequality

In order to investigate the robustness of our income inequality results from the LSMS-dataset, we will make use of the MDHS dataset and its, albeit limited, asset indicators. While the LSMS dataset contained 43 different assets to base the index on, the MDHS dataset contains only 10 indicators, where a number of them have been created from other variables. The limited number of variables decreases the variability of the asset index, and this is the main reason for it not being used in our original analysis. A positive aspect about this index is, however, the high number of observations: 33,798.

Figure 4 displays the wealth distribution of the treatment and control group divided by 2004 and 2010. The Karonga graph shows that the proportion of people in the poorest share is larger in 2004 than in 2010, however, the share of people in the second poorest part is larger in 2010. All in all it is hard to draw any specific conclusions about income inequality based on the distribution in Karonga. In Mzimba and Nkhata Bay, a larger share of people in the second lowest group in 2010 and a larger share of people in the third lowest group in 2004. As with the earlier graph it is hard to draw any particular conclusions based on the graphical inspection.

Figure 4: Distribution of the asset index, for the treatment and control group

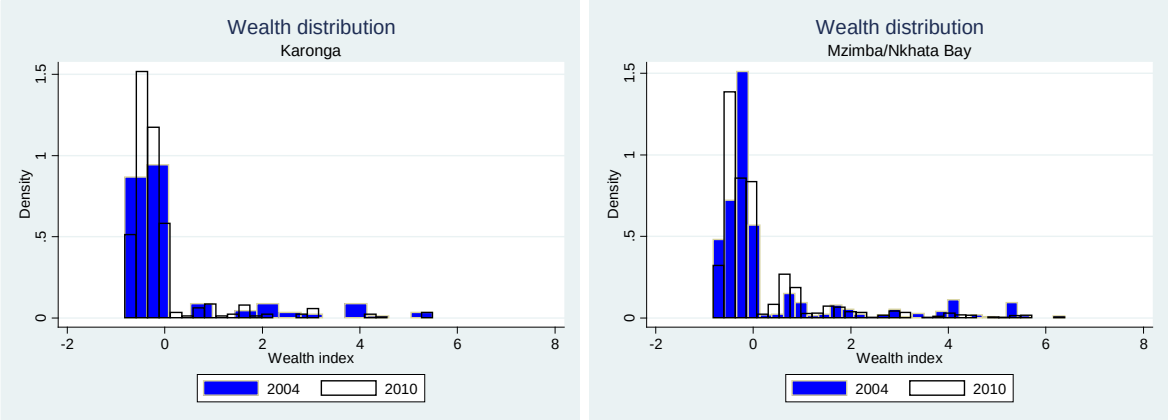


Table 15: Difference in wealth between the 90th and 10th percentile

	Karonga		Mzimba and Nkhata Bay	
Time	2004	2010	2004	2010
Lowest decile	-0.7025	-0.6705	-0.6662	-0.6455
	(0.0192)	(0.0077)	(0.0079)	(0.0053)
Highest decile	2.9212	2.4182	3.0938	2.5034
	(0.2453)	(0.1790)	(0.1153)	(0.0945)
Difference	3.6237	3.0887	3.7600	3.1488
	(0.2288)	(0.0987)	(0.1098)	(0.0658)
Welch t-statistic of difference	-0.5350		-0.6112	
p-value of difference	<0.0001		<0.0001	
Observations	69	264	392	605

We next move to analyse the distance between average wealth in the 10^{th} and 90^{th} wealth percentiles in the treatment and control groups. The results are presented in table 15 and it appears as wealth inequality has decreased in both districts. The change is about 14 percent larger in Mzimba and Nkhata Bay, implying that in Karonga wealth inequality did not decrease as much as in Mzimba and Nkhata Bay during the construction and opening of Kayelekera Uranium Mine. The difference in inequality between Karonga and Mzimba and Nkhata Bay is not as large as with the LSMS data. The difference that was, is now smaller due to the larger decrease in the control group. A slightly worrisome trait is that the average wealth of both the poorest and the richest percentiles has decreased from 2004 to 2010.

Table 16 displays the changes in Gini-coefficients in the treatment and control groups. Similarly to the main LSMS data analysis, the transformed asset index values are used when generating the Gini-coefficients. This does not affect the in-sample distribution, but harms the external validity of the coefficients. Wealth inequality appears to have decreased in both districts. The decrease is from fairly similar levels to a point where Karonga is about 2.5 percent more equal than Mzimba and Nkhata Bay. The decrease is about 45 percent larger in Karonga than in Mzimba and Nkhata Bay.

Table 13: Difference in Gini-coefficients between the treatment and control group

	Karonga		Mzimba/Nkhata Bay	
Time	2004	2010	2004	2010
Gini coefficients	0.5496	0.4982	0.5465	0.5111
	(0.0214)	(0.0169)	(0.0086)	(0.0081)
Difference	-0.0515		-0.0354	
p-value of difference	<0.0001		<0.0001	
Observations	202	749	1118	1678

The results from the MDHS dataset indicate more or less the opposite results than that of the LSMS dataset. This is problematic, specifically as both surveys took place in the same years. The DHS percentile test goes in line with the LSMS test, in that it shows that Karonga has become more unequal during the period. This is however contradicted by the Gini-coefficient test that indicates that Karonga is less unequal than Mzimba and Nkhata Bay. The DHS wealth index is, however, based on a very limited number of variables, wherein most are relatively expensive goods, such as cars, televisions and electricity. We, therefore, believe that the results from the LSMS-analysis are most representative of the truth in this case.

9 Conclusion

In this paper we have investigated the impact of the establishment of the Kayelekera Uranium Mine on HIV incidence and prevalence in the Karonga district in northern Malawi. We performed both qualitative analysis in which we interviewed key informants in Malawi, as well as quantitative analysis using data from MDHS and LSMS.

Our qualitative research is based on interviews performed in and around the city of Karonga and Lilongwe. The principal impacts of the mine are increased economic activity in the area, as well as high levels of immigration to the district. The effect of the mine establishment on HIV incidence is most noticeable through increased consumption of alcohol, fuelled by higher levels of income for mine-workers and others who have benefitted economically from the mine. Alcohol lowers inhibitions and foster risky sexual behaviour. This combined with the cultural expectation of sharing one's income by, for example, treating friends to alcohol, has led to an increase in spread of the HIV. During the construction process there was also significant immigration of sex workers setting up shop in the city and near the mine.

In our quantitative research we analysed HIV and health indicators using the MDHSs of 2004 and 2010. Unfortunately we were not able to apply a difference in differences-estimation using HIV as the dependent variable, as the number of observations of HIV in Karonga is insufficient for statistical inference. We did, however, find that three indicators of risky sexual behaviour changed significantly in Karonga. Since the mine opened the likelihood of a woman being a virgin until marriage, the likelihood of a woman getting pregnant in her teens and the likelihood that a woman's partner is dominant all increased. The first two both have significant impacts on the prevalence of HIV and STIs, whereas the latter only had a significant impact on STI prevalence. The results are somewhat contradictory, as an increase in the number of women being virgins until first union implies a decrease in HIV prevalence and the change in the two other indicators imply increases. As we could not ascertain a clear direction of the change in sexual behaviour, as well as not finding any change in several important indicators, our results indicate that there has been no coherent change in risky sexual behaviour among the women of the Karonga district.

In order to analyse the effects the mine has had on income and wealth we used data from the LSMSs of 2004 and 2010 to create an asset index using multiple correspondence analysis. We then investigated changes in wealth inequality, which, according to literature, is likely to have large impacts on HIV prevalence in Sub-Saharan Africa. The visual representations, differences between 10^{th} and 90^{th} percentiles as well as Gini-coefficients and Lorenz curves, imply that wealth inequality increased more in Karonga following the establishment of the Kayelekera Uranium Mine. The increase in wealth inequality will aid us in understanding the peak in HIV incidence.

Taking all results together, the interviews suggest that risky sexual behaviour has increased as a result of generally higher income levels, which is supported by our findings in the LSMS data. Our results showed that income inequality has increased primarily through increased wealth among the richer part of the population, rather than through increased poverty. The DHS-data on the other hand show no coherent change in risky sexual behaviour. While two indicators signal increases in HIV incidence, one indicator signal decrease and five signal no change.

Among the indicators of risky sexual behaviour that had not changed in Karonga since the mine opened is alcohol consumption. In our interviews, most interviewees argued that

increased alcohol consumption was the main driver of the spread of HIV. It is possible, however, that the alcohol consumption has become more visible in the period since the mine opened, as more boozing dens has opened, moonshining has been banned and cheap sachets of spirit made available. This has changed the composition of alcohol consumption and shifted it into more public areas, where several types of risky sexual behaviour are more likely to occur, for instance transactional and extramarital sex.

Would increased income inequality caused by the mine affect HIV prevalence within the timeframe of this paper? Most likely not, as it takes time for newly employed workers to gather enough assets to be considered wealthy and thereby change the composition of income inequality in the district. The mine commenced production in 2009 and it is the workers who will keep their job for several years from this point, who might accumulate enough income. Hence, the wealth effect of the mine is most likely not visible in the 2010 surveys. Inflation on the other hand appears to have happened rather quickly, as, for instance, the vendors were able to price discriminate by recognizing the workers. We do not, however, see much wealth effect of the inflation in the data for the poorest share of the population. The asset index mean of the poorest percentile has not decreased meaningfully in the LSMS dataset and, in fact, only increased in the DHS dataset. It is rather the positive income shocks that have had an impact on HIV incidence in this time period, as many people temporarily received high incomes. These high incomes were temporary both because many jobs were only available during the construction phase, as well as some people only working short spells due to layoffs.

The positive income shocks experienced by newly employed mineworkers, combined with large inflow of people; the seasonal mineworkers, prospective workers and sex workers, we argue are the main drivers behind the 2009 peak in HIV incidence in the Karonga district. In the two years following the peak we saw a decline in the incidence rate. This we attribute to a number of reasons, for instance, HIV awareness has increased significantly in the last years. Information programs undertaken by NGOs, the Government and Paladin have been successful in educating the public. As a result of this, we often heard in our interviews that people in general know what causes AIDS and how to prevent the virus from infecting them. This result is supported by DHS-data, in which most people have knowledge of the disease. The unfortunate truth here is that the prevalence of boozing makes people disregard the information and contract the disease anyway.

The decline is also affected by people adjusting their view on incomes and them getting accustomed to the high income levels they earn from the mine. Another reason for the decline is that the number of people employed at the mine has decreased in waves in since the mine opened. Ever since the mine opening Paladin has employed more workers than necessary for the daily running of the mining activities. On February 7th 2014, Paladin announced that the Kayelekera Uranium Mine was to be put on care and maintenance. This meant that the work will continue as usual until the current ore stocks are depleted, and then the mine will be inactive until the price of uranium increase. As a result a large share of the workforce would be retrenched and only about 120 to 150 workers will be employed in order to maintain and guard the mine-site. With less people receiving the high mine salaries, the income shock effect will decrease its importance as a driver of HIV and incidence will likely continue to decrease.

What could have been done in order to avoid this unfortunate peak in HIV incidence? With the expectations on the project, we argue that little could have been done other than not going through with it. The investment was the largest in the history of Malawi and people all over the country had grand expectations on the new, up-and-coming mining industry. This inevitably would attract large amounts of people looking for work, especially as Malawi is an extremely poor country and the high salaries at the mine would help anyone feed their families. By international standards, however the Kayelekera Uranium Mine is not that big of an operation. The mine would never be able to supply the inflow of people with jobs. Hence, we argue that the peak in HIV incidence, caused by the positive income shocks and inflow of people, could not have been avoided.

The discussion whether the project should have been done in the first place is not pertinent to this thesis and is best left to others than economists. What ought to be mentioned, however, is that the project has had many positive impacts on the area. The increased income and subsequent multiplier effects have helped many people out of absolute poverty and into relative wealth. As evident of our interviews, a whole new middle-class has emerged in Karonga. For these people the mine has been essential. The city of Karonga has also benefitted from the mine, as the increased economic activity has made the city attract both new businesses and people. The many societal development programs undertaken by Paladin have also positively impacted nearby villages, such as Kayerekera and Bwiliro. Paladin

undertook 70 social development projects between 2007 and 2013. The cost of these projects has been between $12.7 and $16.5 million, which exceed the required $10 million stated in the Development Agreement signed by Paladin and the Malawian government. Therefore it is difficult to claim that the district would have been better off had the project not been undertaken.

The specific traits of Malawi make our results difficult to generalize and apply to different areas. The fact that Malawi is that underdeveloped, without being plagued by war and extreme violence, means that few comparable countries exist. The poverty is a major factor in the analysis, as that many people were attracted to the wages at Kayelekera Uranium Mine.

Another Malawi-specific trait which plays an important role in our analysis is the culture of sharing the wealth and treating your neighbours. The culture of high alcohol consumption has many counterparts throughout the world. This culture of treating your friends, family and neighbours to alcohol as the "*big boss*" in the community is, however, specific to a small number of countries in Sub-Saharan Africa. The positive income shocks provided by working at the mine enabled many poor people to become "*big bosses*" and permitted them and their beneficiaries to engage in the risky sexual behaviour that lead to the peak in HIV incidence. Therefore, our results are best considered for other parts of Malawi, but not other countries with higher incomes and better developed mining industries.

With regards to areas for further research there are several topics that need further investigation. For example, the impact of the various social projects undertaken in the district, and whether they have had a mitigating effect on HIV incidence. It would also be interesting to estimate the impact of the migration to Karonga the mine caused, as we are not able to account for migration in our analysis. Further the follow-up study of the 2010 LSMS survey, which will be published in September 2014, could be used in order to further analyse the effect of the mine in a longer time perspective. It would also be interesting to use this 2013 LSMS as the base year in order to analyse what effect the mine being put on and care and maintenance will have on wealth in the district in the future.

Bibliography

Angrist, J. D. & Pischke, J.-S., 2009. *Mostly harmless econometrics.* New York: Princeton University Press.

Baird, S. J., Garfein, R. S., McIntosh, C. T. & Özler, B., 2012. Effect of a cash transfer programme for schooling on prevalence of HIV and herpes simplex type 2 in Malawi: a cluster randomised trial. *The Lancet,* 379(9823), pp. 1320-1329.

Bertrand, M. & Duflo, E., 2004. How much should we trust differences-in-differences estimates. *The Quarterly Journal of Economics*, 119(1), pp. 249-275.

Burke, M., Gong, E. & Jones, K., 2011. Income Shocks and HIV in Sub-Saharan Africa. *IFPRI discussion papers,* Volume 1146.

Canning, D., 2006. The Economics of HIV/AIDS in Low-Income Countries: The Case for Prevention. *Journal of Economic Perspectives ,* 20(3), pp. 121-142.

Corno, L. & de Walque, D., 2012. Mines, migration and HIV/AIDS in Southern Africa. *Journal of African Economies,* 21(3), pp. 465-498.

de Walque, D. et al., 2012. Incentivising safe sex: a randomised trial of conditional cash transfers for HIV and sexually transmitted infection prevention in rural Tanzania. *BMJ open,* 2(1).

Durevall, D. & Lindskog, A., 2012. Economic Inequality and HIV in Malawi. *World Development, Elsevier,* 40(7), pp. 1435-1451.

Eriksson, P. & Kovalainen, A., 2008. *Qualitative Methods in Business Research.* Los Angeles: Sage.

Fox, A. M., 2010. *The Social Determinants of HIV Serostatus in Sub-Saharan Africa: An Inverse Relationship Between Poverty and HIV?,* Boston: Public Health Reports.

Fritz, K. E. et al., 2002. The association between alcohol use, sexual risk behavior, and HIV infection among men attending beerhalls in Harare, Zimbabwe.. *AIDS and Behavior,* 6(3), pp. 221-228.

Gertler, P., Shah, M. & Bertozzi, S., 2005. Risky business: the market for unprotected commercial sex. *Journal of Political Economy ,* 113(3), pp. 518-550.

Greenacre, M. & Blasius, J., 2006. *Multiple Correspondence Analysis and Related Methods.* London: Chapman & HAll/CRC.

Greenwood, J., Kircher, P., Santos, C. & Tertilt, M., 2013. An equilibrium model of the African HIV/AIDS epidemic. *National Bureau of Economic Research,* Issue No. w18953.

Harttgen, K., Klasen, S. & Vollmer, S., 2013. An African Growth Miracle? Or: What do Asset Indices Tell Us about Trends in Economic Performance?. *Review of Income and Wealth,* 59(S1), pp. S37-S61.

Holmqvist, G., 2009. HIV and Income Inequality: If there is a link, what does it tell us?. *Working paper, International policy Centre for Inclusive growth.*

Holstein , J. & Gubrium , J., 2004. *The active interview.* London: Sage.

Imbens, G. W. & Wooldridge, J. M., 2007. *What's new in econometrics?.* Cambridge: Lecture Notes, NBER Summer Institute 2007.

Kalton, G., 1983. *Introduction to survey sampling.* London: Sage.

Kohler, H.-P. & Tornton, R. L., 2012. Conditional cash transfers and HIV/AIDS prevention: Unconditionally promising?. *The World Bank Economic Review,* 26(2), pp. 165-190.

Kotsadam, A. & Tolonen, A., 2014. African mining, gender and local employment.

Kremer, M. & Morcom, C., 1998. The Effect of Changing Sexual Activity on HIV Prevalence. *Mathematical Biosciences,* 151(1), pp. 99-122.

le Roux Booysen, F. et al., 2007. Trends in povery and inequality in seven African countries. *Cahiers de recherche PMMA,* Issue 6.

Le Roux, B. & Rouanet, H., 2010. *Multiple correspondence analysis.* Vol. 163 ed. s.l.:Sage.

Lechner, M., 2011. The estimation of causal effects by difference-in-difference methods. *SEW Discussion paper no. 2010-28.*

Luke, N., [Unpublished] 2006. *Are wealthy sugar daddies spreading HIV?: exploring economic status, informal exchange, and sexual risk in Kisumu, Kenya.* Los Angeles California, Presented at the Population Association of America 2006 Annual Meeting.

Magruder, J. R., 2011. Marital Shopping and Epidemic AIDS. *Demography*, 48(4), pp. 1401-1428.

Mboup, S., Musonda, R., Mhalu, F. & Essex, M., 2006. HIV/AIDS. In: D. T. Jamison, et al. eds. *Disease and Mortality in Sub-Saharan Africa.* Washington: The World Bank, pp. 237-246.

Morojele, N. K. et al., 2006. Alcohol use and sexual behaviour among risky drinkers and bar and shebeen patrons in Gauteng province, South Africa. *Social Science & Medicine,* pp. 217-227.

Morris, M. & Kretzschmar, M., 1997. Concurrent partnerships and the spread of HIV. *Aids,* 11(5), pp. 641-648.

Oster, E., 2009. HIV and sexual behavior change: Why not Africa?. *National Bureau of Economic Research.*

Oster, E., 2012. Routes of infection: Exports and HIV Incidence in Sub-Saharan Africa. *Journal of the European Economic Assoociation,* 10(5), pp. 1025-1058.

Paladin Energy Ltd., 2012. *Positioned to optimise - Annual Review 2012,* Australia: Paladin Energy Ltd.

Paladin Energy Ltd, n.d. *Kayelekera- Malawi: Backgound.* [Online]
Available at: http://www.paladinenergy.com.au/default.aspx?MenuID=29
[Accessed 02 05 2014].

Parkhurst, J. O., 2010. Understanding the correlations between wealth, poverty and human immunodeficiency virus infection in African countries. *Bulletin of the World Health Organization,* 88(7), pp. 519-526.

Poulin, M., 2007. Sex, money, and premarital partnerships in southern Malawi. *Social Science & Medicine,* 65(11), pp. 2383-2393.

Sahn, D. E. & Stifel, D. C., 2000. Poverty comparisons over time and across countries in Africa. *World development ,* 28(12), pp. 2123-2155.

Swidler, A. & Watkins, S. C., 2007. Ties of Dependence: AIDS and Transactional Sex in Rural Malawi. *Studies in family planning,* 38(3), pp. 147-162.

Tawfik, L. & Watkins, S. C., 2007. Sex in Geneva, sex in Lilongwe, and sex in Balaka. *Social Science & Medicine,* 64(5), pp. 1090-1101.

Watt, M. H. et al., 2012. “Because he has bought for her, he wants to sleep with her”: Alcohol as a currency for sexual exchange in South African drinking venues. *Social Science & Medicine,* pp. 1005-1012.

Wielga, M., Salcito, K. & Wise, E., 2010. *Nomogaia.* [Online]
[Accessed 13 March 2014].

Wilson, N., 2011. Economic Growth and the HIV/AIDS Pandemic: Evidence from the Early 21st-Century Copper Boom. *Williams College, Williamstown.*

Young, A., 2012. The African growth miracle. *National Bureau of Economic Research,* 120(4), pp. 696-739.

Appendix A – Descriptive statistics and Regression results

Appendix Table 1: Summary statistics of the asset index components

Asset index components	Obs	Mean	Std. Dev.	Min	Max
Air conditioner	23550	.0025478	.0504122	0	1
Bed	23551	.3518747	.4775652	0	1
Beer brewing drum	23550	.0379618	.1911079	0	1
Bicycle	23551	.3757802	.484334	0	1
Building items- cement, bricks	23551	.0436075	.2042245	0	1
Car	23550	.0149045	.1211732	0	1
Carpet, rugs, drapes, curtains	23547	.040812	.1978586	0	1
Chair	23551	.418029	.4932455	0	1
Charcoal	23550	.1222505	.3275819	0	1
Cigarettes or other tobacco	23550	.0858599	.280163	0	1
Coffee table (for sitting)	23550	.110276	.3132401	0	1
Clock	23550	.1629299	.3693096	0	1
Cupboard, drawers, bureau	23550	.0648408	.2462499	0	1
Desk	23550	.0055202	.0740941	0	1
Electric or gas stove, hot plate	23551	.0309116	.173082	0	1
Fan	23551	.0303596	.1715785	0	1
Film, film processing, camera	23551	.0075581	.0866098	0	1
Grass for thatching roof	23550	.4822081	.499694	0	1
Insurances	23551	.0078978	.0885196	0	1
Iron (for pressing clothes)	23549	.1746571	.3796816	0	1
Kerosene/paraffin stove	23551	.0135451	.1155949	0	1
Lantern, paraffin	23550	.1986412	.3989857	0	1
Linen- towels, sheets	23551	.3000297	.4582803	0	1
Lorry	23550	.0024204	.0491388	0	1
Mat	23551	.4484311	.4973441	0	1
Matches	23550	.8529087	.3542044	0	1
Mattress	23550	.0263694	.1602347	0	1
Mini-bus	23550	.001741	.0416896	0	1
Mortar/pestle (mtondo)	23549	.4775574	.4995067	0	1
Mosquito net	23551	.1229672	.3284065	0	1
Motorcycle/ scooter	23550	.0048832	.0697107	0	1
Newspapers or magazines	23546	.0370764	.188953	0	1
Paraffin or kerosene	23549	.5948448	.4909325	0	1
Radio (wireless)	23551	.5110186	.4998892	0	1
Refrigerator	23551	.0314636	.1745708	0	1
Sewing machine	23551	.0272175	.1627202	0	1
Sports & hobby equipment	23551	.0064965	.0803406	0	1
Table	23551	.3462273	.4757768	0	1

Tape or CD player, HiFi	23551	.1375313	.3444147	0	1
Television & VCR	23551	.0709099	.2566798	0	1
Upholstered chair, sofa set	23550	.0990658	.2987567	0	1
Washing machine	23551	.0017834	.0421931	0	1
Woodpoles, bamboo	23551	.2652966	.4415004	0	1

Appendix Table 2: Regression on HIV including a number of indicators of risky sexual behaviour

	(1)	(2)	(3)
Dependent variable	HIV	HIV	HIV
Education	0.00813	0.0169	0.00275
	(0.0157)	(0.0176)	(0.0174)
Urban	0.859***		0.800***
	(0.145)		(0.133)
Age-gap	0.0192***	0.0200***	0.0188***
	(0.00716)	(0.00714)	(0.00720)
Virgin-bride	-0.529***	-0.558***	-0.532***
	(0.106)	(0.108)	(0.106)
Alcohol	0.250*	0.252*	0.251*
	(0.129)	(0.135)	(0.132)
Polygamy	0.105	0.0705	0.102
	(0.0751)	(0.0750)	(0.0760)
Teenage pregnancy	0.0754***	0.0720***	0.0756***
	(0.0188)	(0.0191)	(0.0188)
Working	0.118	0.0908	0.119
	(0.0975)	(0.0966)	(0.0978)
Dominance	-0.156	-0.182	-0.154
	(0.109)	(0.111)	(0.109)
Fish	0.115	0.0338	0.120
	(0.255)	(0.245)	(0.256)
Rich		0.452***	0.218
		(0.163)	(0.160)
High medium income		-0.0498	-0.0165
		(0.120)	(0.118)
Low medium income		0.154	0.122
		(0.111)	(0.111)
Poor		0.0136	-0.00117
		(0.112)	(0.113)
Infidelity	-0.131	-0.192	-0.140
	(0.397)	(0.391)	(0.396)
Constant	-2.451***	-2.412***	-2.469***
	(0.211)	(0.218)	(0.230)
Observations	6,695	6,695	6,695

Clustered standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Appendix Table 3: Difference in differences-estimates on four indicators for risky sexual behaviour

Dependent variable	(1) Age-gap	(2) Virgin-bride	(3) Alcohol	(4) Teenage pregnancy
Treatment	1.125***	-3.932***	0.329***	0.0236
	(0.0428)	(0.251)	(0.0815)	(0.137)
Time	0.852	-0.389***	-1.167***	-0.380**
	(1.042)	(0.138)	(0.147)	(0.0830)
Treatment*Time	-2.025	5.123***	-0.0912	0.406*
	(0.894)	(0.163)	(0.140)	(0.123)
Age	0.0488	0.0659***	0.0313***	0.0178
	(0.0285)	(0.0107)	(0.0101)	(0.00628)
Years of education	-0.160	-0.0531	-0.0686	-0.165***
	(0.0601)	(0.0586)	(0.0457)	(0.00766)
Urban	0.353	-2.287***	-0.151	0.00982
	(0.548)	(0.705)	(0.321)	(0.137)
Constant	6.613***	-1.907***	-2.760***	0.00982
	(0.567)	(0.538)	(0.362)	2.236***
Observations	1,174	1,599	3,720	2,033
R-squared	0.011			0.110

Clustered standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Appendix Table 4: T-test of Urban in Karonga

	Obs	Mean	Std. Err.	Std. Dev.	95% Confidence interval	
2004	201	0.1890	0.0277	0.3925	0.1345	0.2437
2010	746	0.1716	0.0138	0.3773	0.1445	0.1987
Difference	947	0.0175	0.0302		-0.0419	0.0768
	t		Degrees of freedom		Pr(|T| > |t|)	
	0.5778		945		0.5636	

Appendix Table 5: Difference in differences-estimates on four indicators for risky sexual behaviour

Dependent variable	(1) STI	(2) STI	(3) STI
Education	-0.0197*	-0.0222*	-0.0230**
	(0.0111)	(0.0115)	(0.0113)
Urban	0.0951		0.0488
	(0.0782)		(0.0867)
Age-gap	-0.0103*	-0.0103**	-0.0104**
	(0.00526)	(0.00525)	(0.00526)
Virgin-bride	-0.337***	-0.338***	-0.337***
	(0.0575)	(0.0569)	(0.0575)
Alcohol	0.222*	0.222*	0.221*
	(0.114)	(0.118)	(0.117)
Polygamy	0.118**	0.115**	0.117**
	(0.0493)	(0.0490)	(0.0493)
Teenage pregnancy	0.0316**	0.0321**	0.0321**
	(0.0133)	(0.0133)	(0.0133)
Working	0.183**	0.182**	0.183**
	(0.0807)	(0.0802)	(0.0804)
Dominance	0.244***	0.245***	0.247***
	(0.0749)	(0.0750)	(0.0748)
Fish	-0.201	-0.204	-0.201
	(0.174)	(0.175)	(0.174)
Rich		0.140	0.127
		(0.0862)	(0.0907)
High medium income		-0.0986	-0.0973
		(0.0770)	(0.0773)
Low medium income		0.0256	0.0237
		(0.0767)	(0.0768)
Poor		-0.0160	-0.0164
		(0.0626)	(0.0625)
Infidelity	0.659***	0.653***	0.654***
	(0.174)	(0.176)	(0.176)
Constant	-2.211***	-2.189***	-2.191***
	(0.114)	(0.122)	(0.123)
Observations	22,864	22,864	22,864

Clustered standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Appendix Table 6: Difference in differences-estimations on risky sexual behaviour

	(1)	(2)	(3)	(4)
Dependent variable	Education	Polygamy	Working	Dominance
Treatment	-0.298	0.293***	-1.069**	-0.0261**
	(0.112)	(0.00359)	(0.513)	(0.00268)
Time	0.433**	-0.332*	-0.524	0.0282
	(0.0918)	(0.175)	(0.860)	(0.0369)
Treatment*Time	-0.0987	0.138	0.512	0.131*
	(0.0914)	(0.167)	(0.859)	(0.0365)
Age	-0.0899***	0.0430***	0.0592***	0.00161
	(0.00665)	(0.00190)	(0.00557)	(0.000718)
Education		-0.0503***	0.0360***	0.00150
		(0.0166)	(0.00592)	(0.00276)
Urban	2.449***	-0.506***	-0.475***	-0.00852
	(0.223)	(0.148)	(0.145)	(0.0473)
Constant	8.682***	-2.124***	-0.819*	-0.0213
	(0.187)	(0.0822)	(0.447)	(0.0267)
Observations	3,720	2,571	3,720	3,720
R-squared	0.180			0.042

Clustered standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Appendix B – Multiple Correspondence Analysis

Multiple correspondence analysis is a manner of analysing underlying structures within nominal categorical data. Multiple correspondence analysis is a generalization of the correspondence analysis of cross-tabulations of two variables to cross-tabulations of multiple variables. Multiple correspondence analysis can be viewed as principal correspondence analysis of categorical variables. It uses all cross-tabulations of a set of variables with themselves and put them in to a squared super-matrix of cross-tabulations, known as a Burt matrix. Alternatively it makes use of ($n \times m$) matrix consisting of n observations of m variables, which will become a Burt matrix if multiplied with its transpose ($Burt = Z^T Z$). On the Burt matrix multiple correspondence analysis will apply the simple correspondence analysis process in which it compare collections of data points' average contributions in order to find an appropriate amount of dimensions that can explain an underlying structure of the data. Simple correspondence analysis calculates proportions of groups of observations which all have the same value of the variables analysed and use these proportions as weights in a calculation of weighted distances. These distances are plotted on a correspondence map and the larger the mass of the weight the closer the geometric centre of the different groups will be. On this correspondence map different dimensions are evaluated in order to see the magnitude in which they explain the relationship between the two variables. These dimensions are then ranked in order of how much explanation ability they carry, or in other words how large proportion of the inertia they explain. The number of dimensions to be included is decided by looking at modified eigenvalues and variance rates. The inclusion of the significant dimensions can be used to create a variable that puts a value on one of the underlying structures in the dataset. In multiple correspondence analysis the simple correspondence analysis algorithm is applied to the Burt matrix and can therefore analyse undercurrents among multiple variables (Greenacre and Blasius, 2006; Le Roux and Rouanet, 2010).

Multiple correspondence analysis is preferred to principal component variable analysis when working with binary variables, such as household possession of specific goods. Where principal component analysis performs orthogonal transformations in order to create correlated variables, correspondence analysis scale the data so that rows and columns are treated equivalently and therefore there is no difference between the explaining and the

explained variables. Similarly correspondence analysis is preferred to factor analysis when dealing with categorical variables as the standard factor analysis assumes continuous variables. Multiple correspondence analysis is also a more sophisticated tool than a plain ranking system in which all variables are added together to create an index.

Appendix C – Questionnaires

Questionnaire used when interviewing workers:

1. Can you see any changes in the district of Karonga, the recent years? (Since the mine opened?)
    - Are there more rich/poor people?
    - Are things more/less expensive nowadays?
    - Have the access to healthcare changed?
    - Have social trust changed?
    - Are people more/less likely to be engaged in stealing?
2. Are there any changes in the alcohol consumption?
    - Are people drinking more/less?
    - What are people drinking?
    - Are people in general more/less violent? Are husbands beating their wives more/less?
3. What has happened to the spread of HIV/AIDS the recent years? (Since the mine opened?)
    - Do people in general know what causes HIV/AIDS?
    - Are people actively trying to protect themselves against HIV/AIDS?
    - Have the access to contraceptives changed? To the better/worse?
    - Is there any stigma/superstition connected to HIV/AIDS?
    - Are there more/less people infected by HIV/AIDS? Is there any difference?
    - Are there any specific groups that have higher HIV rates than others? (young/old, rich/poor, females/males)
4. What are the main drivers behind the HIV/AIDS epidemic according to you?
    - What can be done? What interventions do you believe work?

Questionnaire used when interviewing individuals working in the health sector/NGO`s:

1. What is the direction of HIV incidence in the district of Karonga? (during the construction-phase/production-phase and in the future)
    - Who is getting tested? (rich/poor, high-skilled/low-skilled-workers, males/females)
    - Why are these people getting tested?
    - According to statistics from 2010 the HIV rate was around 11%. Do you believe in this number? Is it realistic?

2. What are the main drivers behind the HIV/AIDS epidemic?
   - Risky sexual behaviour? Alcohol consumption? Economic inequality?
   -Are there any groups in society that are more vulnerable than others? (young/old, rich/poor, females/males)
   - Are there any difference between non-working and working individuals?
   - Are temporary workers likelier or unlikelier to get infected?
   - Are people with shorter life horizons more/less likely to engage in risky sexual behaviour?
   - Is there any change when it comes to violence against women? Does it affect the HIV rate?
3. What has the company Paladin Africa done for the society?
   -Is there any positive/negative effects?
   -Have the company engaged in any social responsibility projects that you know about?

Questionnaire used when interviewing individuals living in the Kayerekera village:

1. What has happened in the village since the establishment of the mine?
   - Can you see any differences in income? (more/less -> rich/poor/gap)
   - What has happened to the price level?
   - Are more people employed now compared to before? (except for the ones working for the mine)
   - Have trust changed in the village?
   - Are there any migrant workers living in the village? If yes, how are they affecting the local population?
   - Is stealing a big problem? (in both the village and at the mine site)
2. Have attitudes regarding HIV/AIDS changed since the mine came?
   - Do people know what causes HIV/AIDS?
   - Are there any stigma/superstition connected to HIV/AIDS?
   - Are more/fewer persons infected today compared to how it was prior to the mine was established?
3. What have the mine company (Paladin Africa) done for the community?

In addition to these three questionnaires, we have also used more specific ones depending on the background of the key informant.