The effect of the national roll-out of anti retro-viral treatment on (HIV-related) mortality in South Africa.

Erasmus University Rotterdam, Erasmus School of Economics

Master's thesis Health Economics

Lotte Beeldman

(431904)

Supervisor: Professor Eddy van Doorslaer

Co-supervisor: Dr. Cobus Burger

Final version: 14-12-2020

Abstract

Background

South Africa is facing one of the largest HIV epidemics in the world and has been since the start of this millennium. Around 2004 the government started the roll-out of a nationwide scheme of free antiretroviral treatment, which treats HIV. Although these medications are proven to be efficacious in medical trials, we estimate the causal impact and effectiveness of this roll-out on (HIV-related) mortality in South Africa.

Data & Method

We have used both district and provincial level data on (HIV-related) mortality, ARV usage, HIV prevalence and poverty. Data sources include Stats SA, the Thembisa Model, the Distict Health Information System and the Multidimensional Poverty Index. We have used a Fixed Effects model in 3 age categories (<15 years old, 15-49 years old and 50+ years old) to estimate the causal effect of ARV usage on annual mortality, in order to control for any unobserved time-invariant heterogeneity between districts and provinces.

Results

Through our estimations, it becomes apparent that ARV usage has had a significant negative effect on overall and HIV-related mortality for those under 50 years old, on both a district level and provincial level. Most importantly, from our provincial analysis we find that a 1 percentage point increase in the ARV usage rate in children under 15 years old decreases their HIV-related mortality rate by 0.318 percentage points. Additionally, we find in our district level analysis that 1 percentage point increase in the overall ARV usage rate decreases the mortality rate in 15 to 49 year old by 0.2 percentage points. In those over 50 years old, the results are more ambiguous.

Conclusions

From our findings, it can be said that the roll-out of free antiretroviral treatment has had a negative effect on the annual (HIV-related) mortality in those under 50 years in particular. Therefore, they suggest that in addition to being efficacious, free provision of ART is also effective. This strengthens the belief that the ART roll-out has had a positive effect on South African's health and should therefore be continued and potentially expanded. This research adds to the existing research on the positive effects of free ART provision in South Africa.

Keywords

HIV-related mortality, South Africa, ARV usage, ART, HIV prevalence, poverty.

Contents

1 Introduction

1.1 HIV/AIDS in South Africa

For over two decades, South Africa has been facing one of the largest HIV epidemics in the world. Its first official case was recorded in 1982, its first AIDS-related deaths in 1985. Already in 1993, it was found that approximately 332.000 people were infected. By 2000, South Africa had the highest HIV prevalence in the world. Around 2006, AIDS-related deaths peak at 210.000 in one year, after which the effects of treatment and prevention of HIV can be observed in lower mortality (UNAIDS, 2020). However, to this day it has proven difficult to control the HIV/AIDS epidemic. Where Uganda has been able to partially turn around the HIV crisis, South Africa still is the leading country where it comes to HIV infections and deaths. In 2018, 7.7 million South Africans were HIV positive, 240.000 got infected and 71.000 died due to HIV/AIDS (Avert, 2020). Bradshaw et al. (2016) have estimated that AIDS has claimed the lives of over 3.1 million South Africans between 1997 and 2012 alone.

In addition to its high mortality, HIV has had a negative impact on all aspects of South African society, on an individual, household, regional and national level. A study by Bachmann and Booysen (2003) found that the baseline income of a household affected by HIV was 0.61 times that of a household unaffected. This indicates that there might be a relationship between poverty and susceptibility to HIV. Additionally, they found that household income declined drastically after diagnosis, exacerbating their poverty. On a macro scale, and more health-related, the United Nations Development Programme (2019) found that life expectancy at birth has fallen from 63.3 years in 1990 to 53.4 years in 2004 and 2005, which was largely caused by HIV (Boutayeb, 2009). It took South Africa 27 years to get back to its mean life expectancy of 1990, which happened in 2017.

From 2002 onward, trials and nationwide schemes have been implemented to provide antiretroviral treatment (ART/ARV) through public healthcare (N.A., 2019). Antiretroviral treatment can counter the effects of HIV on the immune system, so a patient may never develop AIDS. For 2020, South Africa has adopted the 90-90-90 targets, which mean that this year 90% of HIV infected patients must know their status, 90% of patients who know their status must be on ART and 90% of patients on medication must be virally suppressed. This means that 81% of HIV infected South Africans use appropriate medication and that 73% of the HIV infected are virally suppressed. Although by 2018 90% of HIV infected people knew their status, only 62% were on medication and only 54% were virally suppressed (UNAIDS, 2020). These numbers indicate the severity of the situation in South Africa and the length it still has to go in its fight against HIV.

In order to limit the devastating effects of HIV, South Africa has implemented a scheme of free ART, which seems to have a promising effect on overall mortality. This paper aims to further assess the precise effects of ART on South Africa's annual mortality. Where it comes to the impact of ART, it is important to distinguish between clinical efficacy and real-life effectiveness. Clinical

efficacy is tested in a highly controlled setting. The effectiveness assesses the workings of a real-life roll-out and is influenced by factors that are not included in a trial (Bor & Thirumurthy, 2019).

1.2 Clinical efficacy of anti retro-viral treatment

There have been many studies that assess the clinical efficacy of ART. Montaner, Hogg, Raboud, Harrigan, and O'Shaughnessy (1998) state that in 1996 it was found that the now common combination of medications known as ART was the most effective way to treat HIV. Multiple clinical trials showed that this combination could slow down replication of the virus so much that it could no longer be detected (Gulick et al., 1997; Montaner, Reiss, et al., 1998). This is what we now call "virally suppressed". That same year an international panel of physicians recommended this combination as a standard treatment for HIV (Carpenter et al., 1996). Additionally, Wouters, Van Loon, Van Rensburg, and Meulemans (2009) studied the clinical outcomes of ART using a sample of 268 patients. They found that 75% of participants could be classified as treatment successes (CD4 cell count $\geq$ 200 cells/μl and viral load $<$ 400 copies/ml) after 2 years of ART. There are many clinical trials that test the efficacy of ART, all with similar results; ART is proven to be clinically effective.

1.3 ART in practice

Though the efficacy of ART has been proven in a clinical setting, it is much more difficult to assess its effectiveness in a real-life roll-out. Constraints such as inadequate health services, policy failures and/or poor treatment adherence (due to behavioural failures) can greatly impact the effectiveness of a clinically efficacious treatment (Bor & Thirumurthy, 2019). Therefore, to fully understand the real-life impact of ART in South Africa, effects of the roll-out itself must also be evaluated in a setting including all these constraints. Bradshaw et al. (2016) have descriptively studied the annual mortality rates during the ART roll-out and found that the estimated number of AIDS-deaths reached a peak in 2006 (283,000 deaths per year), declined after that and reached 207,000 in 2010. They estimated the annual number of HIV-related deaths using the official disease specific mortality records, but accounted for misclassifications of HIV deaths. This paper shows that there may be an association between the ART roll-out and HIV-specific and overall mortality. More specifically, they have shown that neonatal and infant mortality decline significantly during the ART implementation.

Additionally, Davies et al. (2009) have studied 7 public paediatric ART programmes in the provinces of Gauteng, Western Cape and KwaZulu-Natal. They found that early onset ART in children limits the viral load in young children significantly, improves their overall development and decreases mortality risk. This further strengthens the findings by Bradshaw et al. (2016).

Bor, Herbst, Newell, and Bärnighausen (2013) have studied the effect of ART implementation on life expectancy in KwaZulu-Natal, the province with the highest HIV prevalence in South

Africa. They concluded that between 2003 and 2011 life expectancy has increased by 11.3 years, partly due to the implementation of ART. Also, Tanser, Bärnighausen, Grapsa, Zaidi, and Newell (2013) have researched the effects of ART coverage on the likelihood of becoming HIV-infected in KwaZulu-Natal. They observed 16,667 sero-negative individuals between 2004 and 2011 and found that living in a high ART coverage (30-40%) area decreased the likelihood of becoming sero-postive by 38% compared to living in a low ART coverage (<10%) area. Finally, Burger, van Doorslaer, and Burger (2020) have researched the impact of ART availability on annual mortality, using data from the National Income Dynamics Study (NiDS) in a Fixed Effects regression. They focus on 25-49 year olds, as HIV prevalence is highest in this age group. The research shows that ART availability can cause a decline of up to 30.7% in annual mortality in adults between 25 and 49 years old.

1.4 Design and hypotheses of this study

Although these studies suggest that the ARV roll-out has been effective in diminishing (HIV-related) mortality in South Africa, this research could add to the existing literature. The aim of this paper is to estimate the causal impact of the ART roll-out on overall mortality in South Africa to gain insight into whether or not ART reduces the number of (HIV-related) deaths. Through using different data sources and a different study design, we hope to add to the findings of the previous studies discussed. More specifically, where Bradshaw et al. (2016) did have access to observational data on HIV-related mortality, they did not perform causal estimations. Therefore, we aim to estimate the causal effect of ARV in children under the age of 15 years to further strengthen their findings. Burger et al. (2020) did provide causal estimates, but they did not have access to data on ARV usage. In stead, they used ARV availability, and could therefore not estimate of the effect of the usage itself (including potential lack of treatment adherence) on annual mortality. By estimating the effect of ARV usage itself, we hope to further substantiate and quantify the conclusions provided by Burger et al. (2020).

Although the main purpose of this paper is to add to the existing literature and further quantify the effect of the ARV roll-out on (HIV-related mortality) in South Africa, the outcomes of this study can also have a societal impact. As we are quantifying the effects of the actual ARV roll-out in South Africa, the results will provide an analysis of its effectiveness and could thus be used to potentially adjust or expand the roll-out. Additionally, the results could form a framework for any other countries that may wish to roll-out ART in a similar fashion.

The research question for this study is as follows:

What is the effect of the country wide introduction of antiretroviral treatment on HIV related annual mortality in South Africa?

In order to provide structure to this paper, three hypotheses have been formed based on the existing research discussed earlier. Bradshaw et al. (2016) found a significant decline in neonatal and child HIV-related mortality during the same years as the ART-roll out. We aim to confirm their descriptive findings by estimating the causal effect of the ARV roll-out on mortality in children under 15 years old. The first hypothesis is as follows.

Hypothesis 1: The roll-out of ART in South Africa has had a significant negative effect on annual mortality in children.

Burger et al. (2020) found a large negative effect of ART availability on annual mortality for 25-49 year olds. However, they did not have access to data on ART usage, and could therefore not control for treatment adherence and/or stigma and biases that may inhibit the effects of ART availability. Additionally, they did not have access to HIV-related mortality and performed their analysis on a district level only. Therefore, this paper aims to provide more specific estimates of ART usage on overall mortality for the age group. The second hypothesis of this paper is as follows.

Hypothesis 2: The roll-out of ART in South Africa has had a significant negative effect on annual mortality in the age group of 15 to 49 year olds.

Although some studies find that ART is not as effective in older adults as it is in young ones (Fatti, Mothibi, Meintjes, & Grimwood, 2014), we still expect the ARV roll-out to have a negative effect on overall mortality and HIV-related mortality in those over 50 years old. Therefore, the third hypothesis of this paper is as follows.

Hypothesis 3: The roll-out of ART in South Africa has no significant negative effect on annual mortality in the age group of 50 years or older.

1.5 Cluster levels

Though ART coverage and accessibility is high in South Africa overall, it must be noted that there are large differences between different provinces and districts within provinces. Adam and Johnson (2009) have found that, using the HIV Clinicians Society criteria, 39.9% of eligible patients received ART in the Western Cape in 2008, while in the Free State only 14.2% of eligible patients did. Still in 2017/2018, the Health Systems Trust (2018) stated that while in Limpopo 72% of patients remained on ART, in the North West province this was only 49.1%. Even within provinces, large differences could be found; in Vhembe District (Limpopo) 80.6% remained on ART, when in Capricorn (Limpopo) only 67.2% did. Similarly, large unobserved differences can be expected between districts and even within districts. Therefore, to allow for the most precise estimation of the desired causal effects, we opted for a primary analysis on a district level.

Although precision of estimates is expected to be higher on a district level, we do face limited data availability on this level, as no specific data on HIV prevalence or ART usage is available per age category or gender. Considering the age-component of our hypotheses, this is a considerable downside. Additionally, only overall mortality data is available on a district level, due to new data protection laws regarding sharing and saving of data, currently in place in South Africa. These regulations do not allow HIV-related mortality data to be traced back to individual districts or age groups. However, on a provincial level, data on HIV-related mortality is available, as well as age- and gender-specific data on HIV prevalence and ART usage. Therefore, it was decided to perform a provincial analysis in addition to the district analysis to further solidify our findings on a district level. It must be noted that the provincial level data is mainly modelled data, in stead of observed data. This poses its own set of limitations, that will be discussed in further detail in the data section of this paper.

In the rest of this paper, firstly, the data sets that were used will be introduced including their sources, variable definitions and some important summary statistics. Then the methodology will be discussed, after which the results of the analysis will be presented. Finally, a conclusion and discussion will be included.

2 Data

2.1 District Level Data

2.1.1 Description

On a district level, the following variables were included in the analysis: *Mortality rate, HIV prevalence, ART usage rate* and *poverty headcount*. Table 1 shows the variable definitions. Data is available for the years 2007 to 2015, which means that the ART roll-out initiation (around 2004) will not be covered by this dataset. As was mentioned previously, HIV-related mortality would have been the preferential outcome variable, but no data was available on a district level. However, it is commonly assumed that changes in overall mortality in South Africa are caused by changes in HIV-related mortality. Under this assumption, including the overall mortality rate as a dependent variable should lead to accurate estimates, as we are performing a Fixed Effects analysis that only takes into account changes in the included variables.

The *mortality rate* is based on the total number of deaths in a district level taken from the annual mortality and causes of death reports (Stats SA, 2007b, 2008, 2009, 2010, 2011c, 2012, 2013, 2014, 2015b) and the population data taken from the 2007 and 2016 Community Survey (Stats SA, 2007a, 2016a) and the 2011 Census (Stats SA, 2011a). The population data was interpolated for the remaining years. The data sources do not provide mortality records in the same age groups as we have used, therefore a total mortality has been calculated per age group. *Mortality rate* has

been calculated by dividing the total number of deaths in a specific group by the total population of that same group. This variable is age-specific, but not gender-specific.

ARV usage rates are based on the number of people that are using ARV in a certain district, taken from System (2007-2016), and the total population in that same district taken from the 2011 Census (Stats SA, 2011a) and the 2007 and 2016 Community Survey (Stats SA, 2007a, 2016a). The population data was interpolated for the remaining years. The number of people of ARV in a certain district was determined by summing ARV users per healthcare facility for all facilities in one district. The *ARV usage rate* is calculated by dividing the number of people using ARV by the total number of people in a district. There is no age- or gender-specification in this variable. Therefore, we assume that all age groups and genders face a similar ARV usage rate, which is an inevitable oversimplification of reality, due to data limitations.

Similar to the *ARV usage rates*, we have taken the *HIV prevalences* from the System (2007-2016). Again, HIV positive patients were summed up for all clinics in a district. The HIV prevalence is defined as the total number of HIV positive people in a certain group, divided by the total population of that same group. Similar to the *ARV usage rate*, *HIV prevalence* is neither age- nor gender-specific.

Finally, poverty indicators were taken from the The South African Multidimensional Poverty Index (MPI) (Stats SA, 2011d) for the years 2001 and 2011 as well as the Community Survey Statistical Release of 2016 (Stats SA, 2016b). For all other years the found data was interpolated. Also, as the poverty data from the South African MPI was measured on a municipality level, the weighted average of all municipalities was taken to find the district level poverty rates. This weighted average was based on population data from the 2011 Municipal Report (Stats SA, 2011b) and the 2016 Community Survey Provincial Profiles (Stats SA, 2016c, 2016d, 2016e, 2016f, 2016g, 2016h, 2016i, 2016j, 2016k). Again, for the missing years, the data was interpolated. It must be noted that the Stats SA (2011b) also provides information on population statistics for 2001.

To allow for comparison between the district and provincial analysis, the same age categories will be used in both analyses: $<$15 years old, 15-49 years old and 50$+$ years old. For a full overview of the data availability and its gender- and/or age-specification, table 10 in the Appendix can be consulted.

Table 1: Variable definitions

Variable name	Definition
Mortality rate	Proportion of the population that died.
HIV-related mortality rate	Proportion of the population that died of an HIV-related cause.
ARV rate	Proportion of the population that is using ARV/ART.
HIV prevalence	Proportion of the population that is HIV positive.
Poverty Headcount	Proportion of the population that lives in poverty.

Currently, South Africa has 52 districts, spread out over 9 provinces. However, this was not the case in 2007, and the situation changed several times between 2007 and 2015. Therefore, all data has been attributed to the district they would currently belong to, also for those years where districts may have been different. This is done to ensure consistency in our dataset.

2.1.2 Descriptive Statistics

Table 2 shows the descriptive statistics of the variables used in the district level analysis. As can be seen, the average *mortality rate* across districts and years in those under 15 years old is 0.3%, where in 15 to 49 year olds the average *mortality rate* is 0.8%. In those who are 50 years or older, the average *mortality rate* is 3.36% across districts and years included. ARV rate is 0 in some districts, predominantly in the earlier years included, and goes up to about 0.08. The *HIV prevalence* varies greatly, as can also be seen in the graphs shown later on. The difference between the poorest and richest districts is very large, as in the richest district in the best year only 1.94% of the population lives below the poverty line, while in poorest district in the worst year 33.4% of the population does. This highlights the socioeconomic inequality in South Africa.

Table 2: Summary Statistics District Dataset

	Age	Mean	SD	Min	Max
Mortality rate	<15 years	0.0034	0.0017	0.0003	0.0106
Mortality rate	15-49 years	0.0083	0.0036	0.0006	0.0277
Mortality rate	50+ years	0.0336	0.0086	0.0014	0.0842
ARV rate		0.0078	0.0080	0	0.0801
HIV prevalence		0.1843	0.0779	0.017	0.455
Poverty Headcount		0.1045	0.0653	0.0194	0.334

As can be seen from figure 1, ARV rates in South Africa have increased dramatically between 2007 and 2016, which was to be expected. Over the same period, HIV prevalences have decreased significantly, as can be seen in figure 2.

Additionally, from figure 3 trends in mortality per age category can be observed. This graph clearly shows the large differences in mortality between the age groups. As expected, the 50+

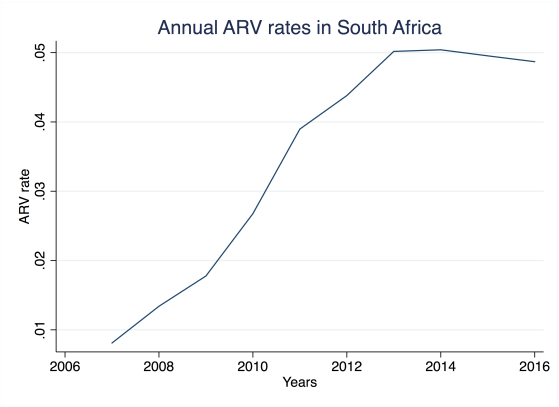

Figure 1: ARV rates in South Africa

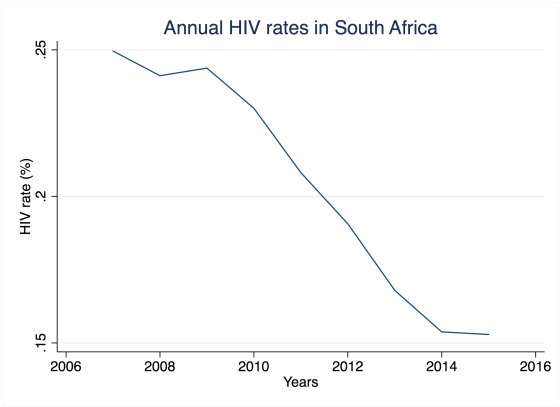

Figure 2: HIV prevalences in South Africa

group faces the largest annual mortality, followed by the group of 15-49 year olds. However, in comparison to this second group, mortality under 15 years old is relatively high. Conclusively, it can be said that all groups see declining mortality over the years, with most noticeable declines in 15-49 year olds. Potentially, these drops could be caused by the effect of ART on mortality earlier found by Burger et al. (2020).

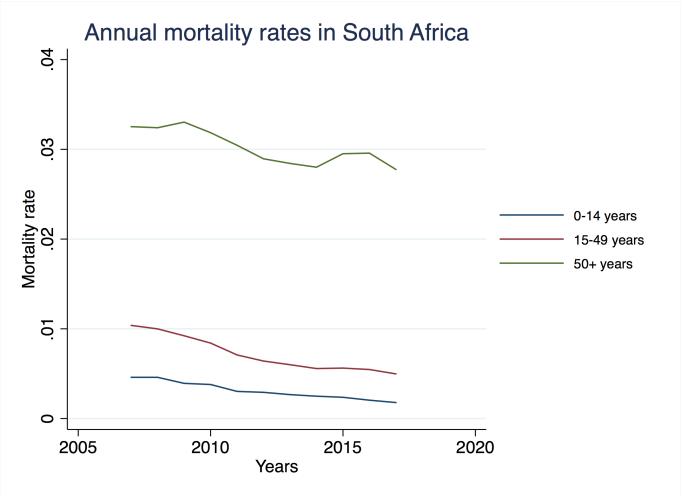

Figure 3: Annual mortality rates in South Africa

2.2 Provincial Level Data

2.2.1 Description

In order to further substantiate the findings of the district level model, a provincial analysis has also been performed. On a provincial level, the following variables were available; *HIV-related mortality rate, HIV prevalence, ART usage* and *poverty headcount*. Table 1 shows the variable definitions. Although this analysis does not allow for the calculation of a detailed effect when it comes to potential heterogeneity within provinces, there are some advantages also. On a provincial

level, data on HIV related mortality is available. This enables the estimation of a more precise effect of ARV usage on mortality, as we know for sure any mortality changes here are HIV related. Additionally, the start of the ART roll-out (around 2004) can be covered by this dataset, as data is available for the years 1997 to 2016. Finally, age-specific data is available on HIV prevalence and ARV usage, which allows for more precise estimates.

The *HIV-related mortality rate* is based on the number of HIV deaths taken from the provincial Thembisa Model (Johnson & Dorrington, 2020b) and population data taken from the mid year population estimates (Stats SA, 2015a). The population data were available for the years 1996, 2001-2011 and 2013-2015. The other years were interpolated. *HIV-related mortality rate* is defined as the number of people that died from HIV related causes in a specific group divided by the total population of that same group. HIV-related mortality data is available for 4 different categories; children under 15 years old, men between 15 and 49 years old, women between 15 and 49 years old and adults over 50 years old. For the middle age group, a gender comparison is therefore also possible. It is important to note that the HIV-related mortality data from the Thembisa Model is modelled data, and is thus not observed. The model takes into account HIV prevalence and ART usage when estimating the HIV-related mortality. Therefore, a large correlation between these independent variables and HIV-related mortality is intrinsic to the data. Although this is a clear downside to using this data, it is the best data source that can be found for HIV-related factors on a provincial level. A full specification of the model and its resulting data set is provided by Thembisa (Johnson & Dorrington, 2020a) and can be consulted for further information.

The *ARV usage rate* is based on the number of people on ARV taken from the Thembisa Model (Johnson & Dorrington, 2020b) and population data taken from the mid year population estimates (Stats SA, 2015a). *ARV usage rate* is defined as the number of people that were using ARV in a specific group divided by the total population of that same group. Data on ARV usage is available for the following categories; children under 15 years old, adults between 15 and 49 years old and adults of 50 years or older. It is important to note that no gender-specific data is available for ARV usage where it is available for the *HIV-related mortality rate*.
HIV prevalence is the proportion of the people that are HIV-positive in a specific group. This data was taken entirely from HIV prevalence data (in percentages) from the Thembisa Model (Johnson & Dorrington, 2020b), but was divided by 100. This data is available in 4 categories; children under 15 years old, men between 15 and 49 years old, women between 15 and 49 years old and adults over 49 years old. Again, similar to the *HIV-related mortality rate*, gender-specific data is available for the middle age group.

Finally, data on *poverty headcount* was taken from the South African MPI (Stats SA, 2011d) for the years 2001 and 2011 as well as the Community Survey Statistical Release of 2016 (Stats SA, 2016b). For the years 2002-2010 and 2012-2015 the data was interpolated. For the years

1997-2000 the data was extrapolated using a report by Hoogeveen and Özler (2005) that estimated poverty headcounts in 1995 and 2000. Although these reports did not estimate the multidimensional poverty index provided by the other data sources, they did allow for the calculation of the relative differences in poverty between 1995 and 2000. They were used to extrapolate the data from the 2001 South African MPI to 1997. No age- or gender-specific information is available on *poverty headcount*.

For a full overview of the data availability and its gender- and/or age-specification, table 10 in the Appendix can be consulted.

2.2.2 Descriptive Statistics

Table 3 shows the descriptive statistics of the provincial dataset. All variables vary greatly over time and between provinces. HIV-related mortality, ARV rate and HIV prevalence have a large standard deviation. However, as this is panel data, some general graphs are more interesting and telling about the data than these descriptive statistics. Therefore, figure 4 and 5 show some general trends in provincial level HIV-related mortality between the years of 1997 and 2015 for men and women respectively, aged 15 to 49 years old.

Table 3: Summary Statistics Provincial Dataset

	Age	Mean	SD	Min	Max
HIV-related mortality rate	<15 years	0.0020	0.0014	0.0001	0.0057
HIV-related mortality rate	15-49 years	0.0048	0.0007	0.0007	0.0126
HIV-related mortality rate	50+ years	0.0027	0.0017	0.0002	0.0065
ARV rate		0.0167	0.0267	0	0.173
HIV prevalence		0.0708	0.0654	0.0018	0.325
Poverty Headcount		0.1266	0.0626	0.0288	0.302

It is clearly visible that there was a steep increase in HIV-related mortality rates for both men and women between 1997 and approximately 2006, after which they declined. These trends suggest that HIV-related mortality may have decreased due to the initiation of the ART roll-out from 2004 onward. For both men and women goes that HIV-related mortality in 2015 was approximately the same in all provinces as it was in 1997. Though the trends of the provinces have converged, still in 2015 a difference in mortality rate of almost 0.002 can be seen between the provinces with low HIV-related mortality (Western Cape and Limpopo) and provinces with high HIV-related mortality (Mpumalanga and Northern Cape) for men. Women between the age of 15 and 49 years experience higher HIV-related mortality than men do, especially around the peak in 2006 in the provinces KwaZulu-Natal and Mpumalanga.

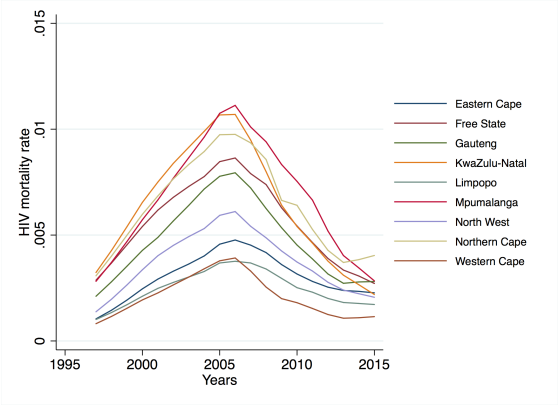

Figure 4: HIV related mortality for men (15-49 years old) per province

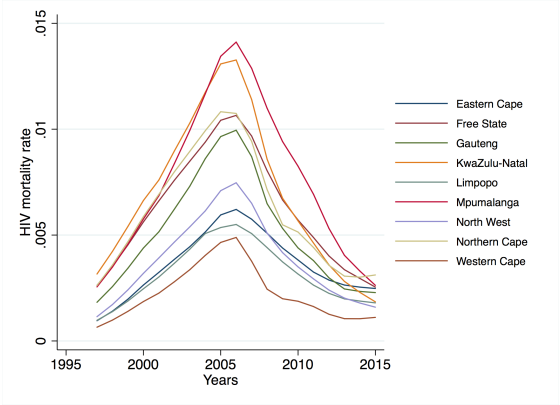

Figure 5: HIV related mortality for women (15-49 years old) per province

Figure 6 shows the provincial ARV rates between 1997 and 2015. The national roll-out of ART was initiated around 2003/2004, which can be seen in the figures. From then on, ART usage grew exponentially until approximately 2012, after which growth stunted slightly though it remained strong. Similar to mortality rates, a difference can be observed between provinces, with KwaZulu-Natal as the province with the highest ARV rates in both men and women. Although this may seem like an advantage, we must be careful to also consider HIV prevalence when looking at the trends in ARV usage in different provinces.

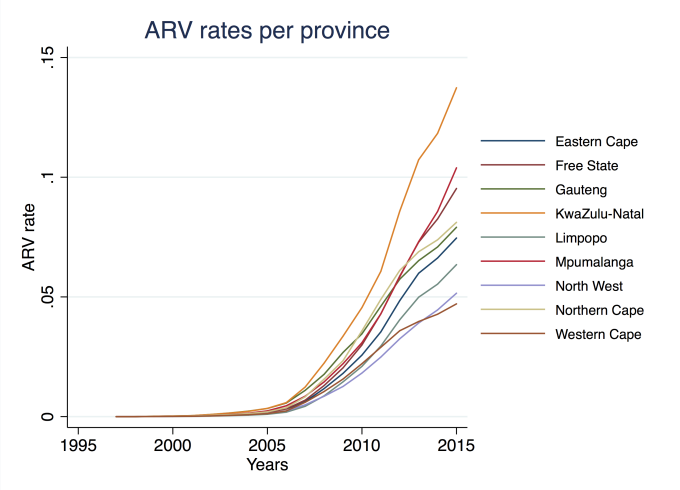

Figure 6: ARV rates per province

Figure 7 and 8 include the general trends in HIV prevalences for both men and women (15-49 years old) between 1997 and 2015. It becomes clear from the graph that women experience much higher rates of HIV prevalence than men do and have been since 1997. Similar to HIV mortality, there is large variation in HIV prevalence between provinces, with Western Cape at the bottom and KwaZulu-Natal at the top of the range. Also, when comparing these results to the trends in ARV

rates discussed previously, it becomes clear that KwaZulu-Natal indeed faces the highest levels of HIV prevalence, which counters the perceived advantage of high ARV usage rates. Finally, it is interesting to see that figure 2 shows a national decline, while our provincial level dataset shows an increase in HIV prevalence over the same period. This could be caused by the fact that our district level data set is based on patient enrollment in clinics, in stead of actual patient calculations. Alternatively, it could be said that the provincial level data is modelled and thus may show an overestimation of HIV prevalence. However, data from UNAIDS (2020) suggests that HIV prevalence in South Africa did actually increase in the period between 2007 and 2015.

Figures 4 to 8 show that the data has been modelled, as all graphs are relatively smooth, which is caused by the assumptions underlying the model.

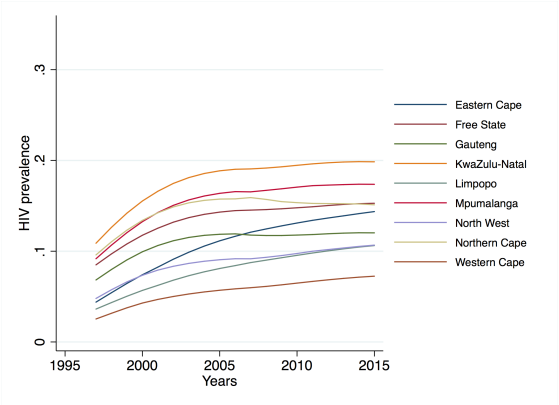

Figure 7: HIV prevalence for men (15-49 years old) per province

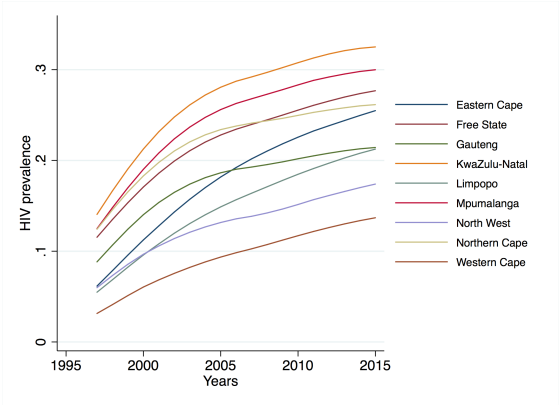

Figure 8: HIV prevalence for women (15-49 years old) per province

Finally, the poverty trends per province are shown in figure 9. Clearly, there are large differences in poverty headcount between the provinces and for all provinces a clear time trend can be observed. Though the poverty headcounts of different provinces have converged, there is still a difference of approximately 10 percentage points between the Eastern and Western Cape. Western Cape has had the lowest poverty headcount since 1999 and also shows the lowest HIV prevalences and HIV related mortality in both genders. However, though Eastern Cape is the province with the highest poverty headcount in all years included, they do not experience the highest HIV prevalences or HIV-related mortality, but are rather in the middle.

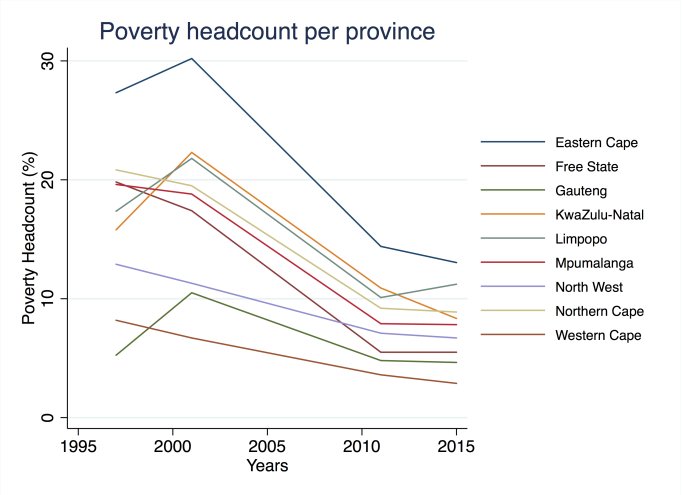

Figure 9: Poverty trends per province

3 Methodology

Considering the effect to be estimated, and the structure of the data, there are several regression types that could be used in this analysis. In similar situations, a pooled ordinary least squares (OLS) regression might be suitable, shown in equation 1.

$$Y_{it} = \beta_0 + \beta_k X_{k,it} + \alpha_i + u_{it} \tag{1}$$

Here, Y_{it} represents the outcome value for individual i at time t, β_0 stands for the constant, X_{it} indicate the control variables over time per individual, α_i is the unobserved heterogeneity and u_{it} is the idiosyncratic error. However, a pooled OLS regression assumes that there is no serial correlation between the error terms, otherwise the pooled OLS estimates might be biased. As we are expecting some serial correlation, in this case a fixed effects (FE) regression is preferred. This type of regression controls for unobserved time-invariant heterogeneity between provinces (and districts) and is commonly used in similar studies. Because provinces (and districts) in South Africa differ significantly where it comes to time-invariant unobserved factors, it is important to account for this heterogeneity. In a FE regression, the time-invariant unobserved individual heterogeneity is controlled for and essentially will be taken out of the equation. Therefore, formula 2 represents the equation for a standard FE regression.

$$Y_{it} = \beta_0 + \beta_k X_{k,it} + u_{it} \tag{2}$$

This formula can be interpreted similarly to the first formula, except for the unobserved heterogeneity that has been taken out. Formula 3 shows the Fixed Effects regression formula that will be used for the district analysis, including all control variables.

$$Mort_{it} = \beta_0 + \beta_1 ART_{it} + \beta_2 HIV_{i(t-1)} + \beta_3 Poverty_{it} + \delta_t + u_{it} \tag{3}$$

In this equation, $Mort_{it}$ stands for the mortality rate in district i in year t. β_0 is the constant and ART_{it} is the ART usage rate per district i at year t. Additionally, a lag of HIV prevalence in district i at year $t-1$ is included as $HIV_{i(t-1)}$. $Poverty_{it}$ represents the poverty headcount of district i at year t. δ_t represents the different time dummies included in the model. Finally, u_{it} indicates the idiosyncratic error. Formula 4 shows the same regression, but including the HIV-related mortality rate for the provincial analysis.

$$HIVmort_{it} = \beta_0 + \beta_1 ART_{it} + \beta_2 HIV_{i(t-1)} + \beta_3 Poverty_{it} + \delta_t + u_{it} \tag{4}$$

It is important to note that an important assumption of a FE regression is that the independent variables are not correlated with the idiosyncratic error u_{it}. This is called the strict exogeneity assumption. If this assumption is violated, estimates may be biased.

In the model, a lag of HIV prevalence has been included, as it is hypothesized that the HIV prevalence of the previous year is more indicative of mortality in the current year than current year HIV prevalence. Current year HIV prevalence is included in the robustness check of the model. Additionally, we will include overall annual mortality in the robustness check instead of HIV related mortality on a provincial level. This is done in order to confirm or decline our hypothesis that HIV-related mortality better represents the effects of ART usage than overall mortality will, even though the Fixed Effects regression only takes into account changes in overall mortality. Additionally, it may mitigate some of the concern we have regarding the intrinsicality of the correlations in the modelled provincial data set.

4 Results

4.1 District Level Analysis

As can be seen from table 4, *ARV usage rate* has a significant effect on overall mortality for both children under 15 years old and adults between the ages of 15 and 49 years old. Both effects are negative, as is to be expected based on the results found by Bradshaw et al. (2016) and Burger et al. (2020). The coefficients found can be interpreted as follows. Firstly, a 1 percentage point increase in the overall ARV usage rate, decreases the mortality rate in children under 15 years old by 0.09 percentage points, ceteris paribus. Although this may seem like a large impact on an average mortality rate of 0.34%, it is important to note that the average ARV rate in the sample is 0.78%, which means a 1 percentage point increase is very significant, and not particularly likely. Previously, no causal effects of ART have been estimated in children under 15 years old. Therefore,

the results we find for this age category cannot be compared to similar papers in order to assess their similarity. However, the findings by Bradshaw et al. (2016) show a decline in mortality rate in children under 15 years, after initiation of the national ART roll-out (approximately 2005/2006). The negative coefficient we find suggests that our findings are in correspondence with the descriptive results found by Bradshaw et al. (2016).

Secondly, a 1 percentage point increase in the overall ARV usage rate decreases the mortality rate in 15 to 49 year olds by 0.2 percentage points, ceteris paribus. Again, on an average mortality rate of 0.83%, this seems a very large effect. If, however, we estimate the effect of a 0.1 percentage point increase, which seems much more likely on an average of 0.78%, we find that this results in a 0.02 percentage points decrease in the mortality rate of 15 to 49 year olds. This estimate seems plausible. Although these results cannot be directly compared to those by Burger et al. (2020), they do also find a negative impact of the ART roll-out on mortality in adults between 15 and 49 years old. However, they included ART availability as an independent variable, where we used ART usage.

For adults over the age of 50 years old, we find a positive effect that is not significant at a 10% level. A positive effect of ARV usage on mortality is unexpected and cannot be easily explained. We hypothesize that this effect is caused by a possible decline in other healthcare resources (more useful to the elderly) that occurred simultaneously with the ARV roll-out (Maphumulo & Bhengu, 2019). This would mean that the positive effect of the ARV usage rate on mortality measured here is not caused by the ARV usage rate, but by an alternative cause. However, we do not have any estimations that could quantify this effect. Also, some studies show that ART should be age-appropriate and is not as effective in those over 55 years old as it is in younger adults (Fatti et al., 2014). This is another potential explanation for the positive effect found.

In children and adults between 15 and 49 years old, a negative time trend is predicted, for adults over 50 a positive time trend is found. As is to be expected, the lagged HIV prevalence has a positive impact on mortality in all age categories. However, only in children under 15 and adults between 15 and 49 years old this effect is significant at a 10% significance level. Finally a negative effect of poverty headcount is found in children under 15 years old, which is peculiar. Intuitively, it seems likely that a higher poverty headcount would lead to higher overall mortality. However, as can be seen in the table, the negative effect found is very small, and insignificant on a 10% significance level. For the older age groups a positive effect is found, but those are also insignificant at a 10% level. Therefore, we may be able to conclude that poverty headcount has no significant effect on mortality. A potential explanation is that the trend in poverty is included in the time dummies.

It is important to note that the effect of the ARV usage rate calculated in this district level

analysis, is the effect of changes in the overall ARV usage rate of the population of a district, and does thus not measure the effect of changes in the ARV usage rate of the specific age group itself. Due to data limitations, we have had to assume that the ARV usage rate is equal in all age groups, which clearly is an oversimplification of the actual situation. The same applies to the HIV prevalence. The time dummies have not been included in table 4. A full table including the time dummies can be found in table 11 in the Appendix.

Table 4: District FE regressions per age category

	(1)	(2)	(3)
	Mortality rate	Mortality rate	Mortality rate
	< 15 years old	15-49 years old	50+ years old
ARV rate	-0.0938**	-0.241***	0.243
	(0.0394)	(0.0791)	(0.267)
Lag HIV rate	0.00456**	0.0112**	0.0275*
	(0.00199)	(0.00520)	(0.0150)
Poverty headcount	0.000406	0.0119**	0.00637
	(0.00329)	(0.00590)	(0.0249)
Constant	0.00409***	0.00762***	0.0262***
	(0.000568)	(0.00146)	(0.00461)
Observations	463	463	463
R-squared	0.616	0.618	0.087
Number of groups	52	52	52

Robust standard errors in parentheses

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

As there is no district level information per gender available, a separation between men and women is not possible. On a provincial level, this separation will be made.

4.2 Provincial Analysis

As was previously mentioned, on a provincial level data was available for the age groups <15 years old, 15-49 years old and 50+ years old. Therefore, these groups were adhered to in both analyses to allow for comparison. Additionally, for the age group of 15-49 years old, the analysis was performed for men and women separately. The results of the regressions for the different age groups are presented in table 5.

A significant negative effect of the ARV usage rate on HIV-specific mortality was found for all age categories. In children under 15 years old a 1 percentage point increase in their ARV usage rate results in a 0.318 percentage points decrease in their HIV-related mortality rate, ceteris paribus. This effect seems very large, taking into account the average HIV-related mortality rate in those

under 15 years old, but is important to note that the average ARV usage rate amongst those under 15 years old is 0.3%, which means that a 1 percentage point increase would mean a quadrupling of the average ARV usage rate. Even a 0.1 percentage point increase would be relatively large compared to its average. Therefore, it seems as though this estimate is reasonable. In 15-49 year olds a 1 percentage point increase in ARV usage rate decreases their HIV-related mortality rate by 0.0734 percentage points, ceteris paribus. For people aged 50 or older goes that a 1 percentage point increase in the ARV usage rate causes a 0.0195 percentage point decrease in HIV-related related mortality, ceteris paribus.

The large effect of the ARV usage rate found in children under 15 years of age is quite surprising as most often the largest effects are observed in the category of 15-49 year olds. This is also what Burger et al. (2020) found in an analysis similar to ours. However, they did not have access to HIV related mortality and ARV usage, which can be of great importance for the outcomes found. This will be further discussed in our comparison of the results.

The lagged HIV prevalence has a significantly positive effect on HIV-specific mortality rate for all age categories and the poverty headcount has a significantly positive effect in the higher two age categories. Both of these effects are expected. A 1 percentage point increase in the HIV prevalence in children under 15 years old, increases the HIV mortality rate in that age category by 0.058 percentage points the next year. For 15-49 year olds an HIV prevalence increase of 1 percentage point increases the HIV mortality rate of the next year by 0.054 percentage points and for those over 50 a 1 percentage point increase in the HIV prevalence increases HIV mortality by 0.064 percentage points in the next year.

Finally, only in those over the age of 50 years old a significant negative effect of poverty headcount is found. As previously mentioned in the district analysis, a negative effect of poverty on mortality seems counter-intuitive. Although the found effect is significant, it is very small.

Time dummies were included in the analysis, but not in table 5. The full results can be found in table 12 in the Appendix.

Table 5: Provincial FE regressions per age category

	(1)	(2)	(3)
	HIV mortality rate	HIV mortality rate	HIV mortality rate
	$<$ 15 years old	15-49 years old	$>$ 50 years old
ARV rate	-0.318***	-0.0734***	-0.0195***
	(0.0289)	(0.00583)	(0.00401)
Lag HIV prevalence	0.0579***	0.0542***	0.0637***
	(0.0129)	(0.0114)	(0.0142)
Poverty headcount	0.00349	0.00675	-0.00825**
	(0.00362)	(0.00508)	(0.00306)
Constant	0.000930	-0.00274	0.00124*
	(0.000675)	(0.00163)	(0.000647)
Observations	162	162	162
R-squared	0.950	0.945	0.866
Number of groups	9	9	9

Robust standard errors in parentheses

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

To further assess the effects of ART usage on adults between 15 and 49 years old for the genders separately, the same analysis has been performed for both men and women. The results of these regressions are presented in table 6. As can be seen, the effect of ARV usage on the HIV mortality rate is significantly negative for both men and women, but is larger for women than for men. In men between 15 and 49 years old, a 1 percentage point increase in the use of ARV, decreases the HIV mortality rate in that same category by 0.061 percentage point. In women between 15 and 49 years old, a 1 percentage point increase in ARV usage, decreases the HIV mortality rate by 0.085 percentage points. The effect of the HIV prevalence of the previous year is larger for men than it is for women. In men, a 1 percentage point rise in the HIV prevalence, will increase the HIV related mortality rate by 0.060 percentage points the following year. For women, a 1 percentage point increase in the HIV prevalence will lead to an increase in the HIV related mortality rate of 0.052 percentage points. Though all these effects are significant, they are not particularly large. No significant effect of poverty headcount was found in either gender.

The results including time dummies, can be found in table 13 in the Appendix.

Table 6: Provincial FE regressions per gender

	(1)	(2)
	HIV mortality rate	HIV mortality rate
	Men 15-49 years old	Women 15-49 years old
ARV rate	-0.0606***	-0.0846***
	(0.00515)	(0.00794)
Lag HIV prevalence	0.0603***	0.0517***
	(0.0116)	(0.0111)
Headcount	0.00369	0.0101
	(0.00439)	(0.00567)
Constant	-0.00194	-0.00375*
	(0.00127)	(0.00189)
Observations	162	162
R-squared	0.941	0.944
Number of groups	9	9

Robust standard errors in parentheses

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

4.3 Comparison district and provincial estimates

There are some important differences between the district and provincial level estimates that must be discussed. Most importantly, we find an extraordinarily large effect of ARV usage rate on HIV-related mortality in children under 15 years old in the provincial model. We also find a negative, though much smaller effect in the district level analysis. Compared to the results by Burger et al. (2020), where the largest effects were found in adults between 25 and 49 years old, our provincial level estimates are quite unexpected. One potential explanation for the large difference between our district level (expected) estimates and our provincial estimates is that on a provincial level we measure the effect of an increase in the ARV usage rate in children under 15 years old only. On a district level we measure the effect of an increase in the overall ARV usage rate of the population. Therefore, it can be argued that the provincial level estimates in children are much larger, because an increase in the overall ARV usage rate might often be disproportionately concentrated amongst adults. An additional, but less likely explanation could be that ARV usage rates have a larger effect on changes in HIV-related mortality than they do on changes in overall mortality, particularly in children under 15 years old. The underlying assumption of the district level estimates is that changes is overall mortality are caused by changes in HIV-related mortality in all age groups, but this might not be the case for young children. However, the following robustness analysis shows that only a small part of the difference can be explained by this reasoning.

However, this does not explain the large difference between our district level and provincial level estimates in the adult (15-49 years old) category. Under the explanation above, we would

expect the effects of ARV usage rate to be similar on a district and provincial level in this age category, which is not the case. The large discrepancy between the two estimates might be caused by increased heterogeneity on a provincial level.

4.4 District Robustness Analysis

In order to further substantiate our results, we have performed several robustness checks. On a district level, we tested if our assumption that lagged HIV prevalence was of greater importance to the overall mortality than current year HIV prevalence. From our results in table 7, it becomes apparent that most estimates remain the same. However, it does lead to insignificance of the coefficient for the ARV usage rate in children under 15 years old. Also, the resulting change in coefficient of 0.02 is rather large. This could indicate that it matters whether we include HIV prevalence or its lag. Potentially, this means that when we include HIV prevalence, this measures an effect similar to that of the ARV usage rate. This would mean that including a lag of HIV prevalence is actually preferred to including current year HIV prevalence.

Table 7: District robustness regressions using HIV prevalence per age category

	(1)	(2)	(3)
	HIV mortality rate	HIV mortality rate	HIV mortality rate
	<15 years old	15-49 years old	50+ years old
ARV rate	-0.0734	-0.256***	0.243
	(0.0468)	(0.0858)	(0.290)
HIV rate	0.00491*	0.0107*	0.0228
	(0.00247)	(0.00550)	(0.0142)
Poverty headcount	-0.000381	0.0142	0.0184
	(0.00555)	(0.0103)	(0.0368)
Constant	0.00406***	0.00784***	0.0265***
	(0.000935)	(0.00165)	(0.00544)
Observations	464	464	464
R-squared	0.594	0.617	0.071
Number of groups	52	52	52

Robust standard errors in parentheses

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

4.5 Provincial Robustness Analysis

The assumptions underlying the provincial analysis have also been checked in a robustness analysis. As in the district analysis, the HIV prevalence of the current year was included in the model in stead of the lagged variable to see whether our assumption holds that lagged HIV prevalence is of greater importance that current HIV prevalence. The results are presented in table

8. As can be seen from this table, the results are largely robust to using HIV prevalence. However, the effect of HIV prevalence on HIV-related mortality is somewhat larger and more significant than the effect of lagged HIV prevalence. This could indicate that HIV prevalence is of greater importance than lagged HIV prevalence. However, the effect of including HIV prevalence instead of lagged HIV prevalence does not change the previously found large effect of the ARV usage rate on HIV-related mortality in children. A similar analysis has been performed separately for the two genders, see table 14 in the Appendix. The results are similar to the robustness analysis of age groups. We can therefore conclude that although our assumption may not be correct, this is not of great importance in this research.

Table 8: Provincial robustness regressions using HIV prevalence per age category

	(1)	(2)	(3)
	HIV mortality rate	HIV mortality rate	HIV mortality rate
	<15 years old	15-49 years old	50+ years old
ARV rate	-0.311***	-0.0677***	-0.0212***
	(0.0281)	(0.00566)	(0.00401)
HIV prevalence	0.0810***	0.0487***	0.0745***
	(0.0141)	(0.0121)	(0.0124)
Poverty headcount	0.00360	0.00317	-0.00481
	(0.00291)	(0.00503)	(0.00271)
Constant	0.000443	-0.00239	0.000185
	(0.000562)	(0.00158)	(0.000530)
Observations	171	171	171
R-squared	0.956	0.937	0.886
Number of groups	9	9	9

Robust standard errors in parentheses

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

As a second robustness test, overall mortality was included in the analysis rather than HIV related mortality. It was assumed that HIV-related mortality is a better dependent variable to use than overall mortality, which is often used in similar analyses. However, HIV-related mortality data in only available from the Johnson and Dorrington (2020b), which included modelled rather than observed data. Therefore, we are uncertain whether the significantly negative effects we found in the provincial analysis are accurate estimates of the effect of ARV, or are instead reflections of the modelling process behind the data. By performing the same analysis on overall mortality in stead of HIV-related mortality, the robustness of our findings can be determined. The robustness results for the age groups can be found in table 9. They show that the results are relatively similar, except for those over 50 years old. The unexpectedly large effect in children under 15 years old remains, which could indicate that our results are fairly robust. Although observed data

remains preferential over modelled date, we find that modelling of data does not intrinsically lead to unreliable results in this case.

Table 9: Provincial robustness regressions overall mortality per age category

	(1) Mortality Rate <15 years old	(2) Mortality Rate 15-49 years old	(3) Mortality Rate 50+ years old
ARV rate	-0.300**	-0.0739***	-0.0515***
	(0.0909)	(0.00922)	(0.0142)
Lag HIV prevalence	0.0736	0.0430	0.0863
	(0.0538)	(0.0249)	(0.0837)
Poverty headcount	-0.00852	0.00542	-0.0904***
	(0.00608)	(0.00939)	(0.0234)
Constant	0.00489***	0.00193	0.0472***
	(0.00119)	(0.00313)	(0.00509)
Observations	162	162	162
R-squared	0.744	0.916	0.643
Number of groups	9	9	9

Robust standard errors in parentheses

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

Similarly, overall mortality was used in the gender level analysis as well. The results are presented in table 15 in the Appendix. Expectantly, inclusion of overall mortality rather than HIV-related mortality as a dependent variable leads to similar results in both men and women between 15 and 49 years old. Especially in this age category, it seems as though inclusion of modelled HIV-related mortality data does not lead to biased results.

5 Conclusions

The conclusions that can be drawn from our findings will be discussed per hypothesis. The first hypothesis states that the ARV rate has a significantly negative effect on the mortality rate in children under 15 years old. This hypothesis can be confirmed based on both the district and provincial analysis. In our district analysis a negative effect of -0.0938 was found for children, while in the provincial analysis a negative effect of -0.318 was found. Although the effect in the provincial analysis was much larger, this can be explained by the different measures of ARV usage rates. These results are in correspondence with previous literature on the topic (Bradshaw et al., 2016; Kerber et al., 2013; Ndirangu, Newell, Tanser, Herbst, & Bland, 2010).

The second hypothesis of this paper says that there is negative significant effect of ARV usage on mortality in 15-49 year olds. Again, this hypothesis is confirmed by both the district and

provincial analysis. On a district level a negative effect of -0.241 was found, on a provincial level the effect was smaller at -0.0734. Both effects are significant at a 1% significance level. The results are in line with the findings of Burger et al. (2020).

Finally, the third hypothesis says that there is also a significantly negative impact of ARV usage on mortality in adults of 50 years or older. The results on this hypothesis are more conflicted. Though we find a small negative, but significant (at 1%), effect in the provincial model, we find a large (0.243) insignificant positive effect in the district analysis. The provincial results are as expected, considering the relatively low HIV prevalence rates in the population above 50 years old. However, the district findings are not as easily explained. We propose that it may be caused by a lack of investments in elderly care, due to a bigger investment in the ARV roll-out (Maphumulo & Bhengu, 2019).

Comparing both results, we find particularly large differences in the estimates of the child and adult age categories. However, they can be explained by different underlying assumptions about specificity of ARV usage rates, (HIV-related) mortality and heterogeneity of the cluster (provinces vs. districts).

We found that our results are mostly robust to the inclusion of current year HIV prevalences. When including overall mortality rathen than HIV-related mortality as our dependent variable, we find similar results in both children under 15 and adults between 15 and 49 years old. Therefore, we conclude that although observed data remains preferential, the modelling of data does not automatically lead to unreliable results.

Conclusively, the answer to the research question of this paper is that the nationwide ARV roll-out has had a significantly negative effect on (HIV-related) mortality in South Africa in those under 50 years old. In older people, the results are more conflicted and ambiguous.

6 Discussion and Recommendations

This research makes use of nationwide data on ART usage and HIV related mortality on a district and provincial level to estimate the causal impact of the national ART roll-out on annual mortality in South Africa. Though similar studies have been performed, ours is the first to estimate the causal impact based on actual ART usage, rather than ART availability in the area, and HIV-related mortality in addition to overall mortality.

We have found that ART usage has a negative impact of annual (HIV-related) mortality in those under 50 years old. In those over 50, the results are slightly more ambiguous. Also, for both genders in the 15-49 year olds a negative effect was found. The most striking finding is the very large negative impact found in children on a provincial level, as in other similar studies the largest

effect of the ART roll-out was found in 15-49 year olds. Although our results are very promising, there are some limitations to this research.

6.1 Data Limitations

First of all, there are some data limitations that must be considered. The limitations will be discussed independently for the district and provincial analysis. On a district level, the dataset used is fairly limited. No age specific data is available on HIV prevalence and ARV usage. This could potentially influence the accuracy of the results, as the data used is very general. Additionally, no gender specific information was available on any of the indicators used, which similarly may influence the precision of the estimates, although expectantly to a lesser extent. Finally, it is important to indicate that no HIV-related mortality data is available on a district level for the years included. Therefore, this analysis is based on the assumption that changes in overall mortality for the years included are mostly due to changes in HIV-related mortality. Consequently, it can be argued that using overall mortality in the district analysis does not lead to imprecision. However, if this assumption turns out to be false, this means that the results may be inaccurate.

Also, on a provincial level there are some data limitations. The most important limitation is the data source. Except for the poverty headcount, all data was taken from the provincial Thembisa Model (Johnson & Dorrington, 2020b). This is modelled data that is not primarily based on actual observations, and thus maybe inconsistent or biased. As previously discussed, when taking a closer look at the model, it can be seen that the HIV-mortality data that is used is modelled using the HIV prevalence and ARV availability in a certain area. Therefore, it can be assumed that there is significant and intrinsic correlation between the independent variables and the HIV-related mortality as HIV-related mortality is partly modelled on these independent variables. It is therefore not surprising that we find such strong correlations in our own analysis. Although this may be a cause of bias, it must also be noted that when using overall mortality as a dependent variable, in stead of HIV-related mortality, similar results are found. This could indicate that our results are fairly robust. However, observed HIV-related mortality data would improve the solidity of our results.

Additionally, as the age groups from the Thembisa model are fairly large, neonates could not be included in this analysis separately, which would have been interesting as Bradshaw et al. (2016) found an association that suggests a potential large negative effect of ART usage on neonatal mortality. Finally, data on ART usage and HIV prevalence could only be found for genders separately in the adult age category (15-49 years old). If more data had been available a more precise analysis could have been performed.

6.2 Limitations Fixed Effects Regression

Considering the model of choice (a Fixed Effects Regression), some limitations can also be identified. It is important to note that this model only controls for the included time variant independent variables (*HIV prevalence* and *poverty headcount*), time-invariant unobserved heterogeneity and the included time dummies. It can therefore not control for time-variant unobserved heterogeneity. That means that any variance between districts and/or provinces that is not constant over time, is not controlled for in this analysis. This could potentially lead to bias, as we do not know for sure that there are no time-variant omitted variables. Especially considering our extensive data limitations this should not be overlooked.

6.3 Recommendations

Considering the findings from this paper and its limitations, we have several recommendations for future research and data availability. First of all, it is important to note that this research could be improved by using more detailed and precise data on HIV prevalence, ART usage and HIV-related mortality. For example, gender- or age-specific data on these factors would allow for a more precise estimation of the causal effects provided in this paper. Currently, this data is not available on either a provincial or district level. Therefore, our first recommendation would be to allow for the publication of this specific data to improve future study results.

Considering the findings from this paper, further research should focus on the effects of ART usage on HIV-related mortality in children under 15 years old in particular. The use of more sophisticated data sets could potentially improve the evidence provided in this paper that shows a negative effect of ART usage on HIV-related mortality in this age category, especially on a provincial level. This goes back to the earlier mentioned data limitations that were faced in this research. It would be even more thorough to include age- and gender-specific HIV-related mortality and ARV usage data in the district level analysis, to allow for the use of smaller, less heterogeneous clusters. The findings on effectiveness of ARV in adults are mostly in line with previous research and do therefore not need as much expansion, though better data would improve their solidity.

The findings of this paper suggest that the national roll-out of ARV has improved the health of the South African population. Therefore, when it comes to policy recommendations, they suggest that the roll-out of free ARV has been at least partially effective and should therefore be upheld. Finally, they show other countries facing similar HIV epidemics that free ARV does have an effect on mortality, and is thus a solid policy solution.

7 Bibliography

Adam, M. A., & Johnson, L. F. (2009). Estimation of adult antiretroviral treatment coverage in south africa. *South African Medical Journal*, *99*(9).

Avert. (2020). *Hiv and aids in south africa.* https://www.avert.org/professionals/hiv-around-world/sub-saharan-africa/south-africa, note = "Accessed: 2020-03-14".

Bachmann, M. O., & Booysen, F. L. (2003). Health and economic impact of hiv/aids on south african households: a cohort study. *Bmc public health*, *3*(1), 14.

Bor, J., Herbst, A. J., Newell, M.-L., & Bärnighausen, T. (2013). Increases in adult life expectancy in rural south africa: valuing the scale-up of hiv treatment. *Science*, *339*(6122), 961–965.

Bor, J., & Thirumurthy, H. (2019). Bridging the efficacy–effectiveness gap in hiv programs: Lessons from economics. *JAIDS Journal of Acquired Immune Deficiency Syndromes*, *82*, S183–S191.

Boutayeb, A. (2009). The impact of hiv/aids on human development in african countries. *BMC Public Health*, *9*(S1), S3.

Bradshaw, D., Msemburi, W., Dorrington, R., Pillay-van Wyk, V., Laubscher, R., & Groenewald, P. (2016). Hiv/aids in south africa: how many people died from the disease between 1997 and 2010? *Aids*, *30*(5), 771–778.

Burger, C., van Doorslaer, E., & Burger, R. (2020). The health impact of free access to antiretroviral therapy in south africa. *The Lancet*. https://papers-ssrn-com.eur.idm.oclc.org/sol3/papers.cfm?abstract_id=3506209.

Carpenter, C. C., Fischl, M. A., Hammer, S. M., Hirsch, M. S., Jacobsen, D. M., Katzenstein, D. A., ... others (1996). Antiretroviral therapy for hiv infection in 1996: recommendations of an international panel. *Jama*, *276*(2), 146–154.

Davies, M.-A., Keiser, O., Eley, B., Rabie, H., van Cutsem, G., Giddy, J., ... Moultrie, H. (2009). Outcomes of the south african national antiretroviral treatment programme for children: the iedea southern africa collaboration. *South African medical journal*, *99*(10).

Fatti, G., Mothibi, E., Meintjes, G., & Grimwood, A. (2014). Antiretroviral treatment outcomes amongst older adults in a large multicentre cohort in south africa. *PloS one*, *9*(6), e100273.

Gulick, R. M., Mellors, J. W., Havlir, D., Eron, J. J., Gonzalez, C., McMahon, D., ... others (1997). Treatment with indinavir, zidovudine, and lamivudine in adults with human immunodeficiency virus infection and prior antiretroviral therapy. *New England Journal of Medicine*, *337*(11), 734–739.

Health Systems Trust. (2018). *District health barometer 2017/2018.* https://www.hst.org.za/publications/Pages/DHB20172018.aspx.

Hoogeveen, J. G., & Özler, B. (2005). Not separate, not equal: Poverty and inequality in post-apartheid south africa.

Johnson, L., & Dorrington, R. (2020a). *Report on the provincial hiv estimates.* https://www.thembisa.org/downloads.

Johnson, L., & Dorrington, R. (2020b). *Thembisa model version 4.3.* https://www.thembisa.org/

downloads.

Kerber, K. J., Lawn, J. E., Johnson, L. F., Mahy, M., Dorrington, R. E., Phillips, H., ... others (2013). South african child deaths 1990–2011: have hiv services reversed the trend enough to meet millennium development goal 4? *AIDS (London, England)*, *27*(16), 2637.

Maphumulo, W. T., & Bhengu, B. R. (2019). Challenges of quality improvement in the healthcare of south africa post-apartheid: A critical review. *Curationis*, *42*(1), 1–9.

Montaner, J. S., Hogg, R., Raboud, J., Harrigan, R., & O'Shaughnessy, M. (1998). Antiretroviral treatment in 1998. *The Lancet*, *352*(9144), 1919–1922.

Montaner, J. S., Reiss, P., Cooper, D., Vella, S., Harris, M., Conway, B., ... others (1998). A randomized, double-blind trial comparing combinations of nevirapine, didanosine, and zidovudine for hiv-infected patients: the incas trial. *Jama*, *279*(12), 930–937.

N.A. (2019). *Hiv aids in south africa timeline 1940s-2009.* https://www.sahistory.org.za/article/hivaids-south-africa-timeline-1940s-2009.

Ndirangu, J., Newell, M.-L., Tanser, F., Herbst, A. J., & Bland, R. (2010). Decline in early life mortality in a high hiv prevalence rural area of south africa: evidence of hiv prevention or treatment impact? *AIDS (London, England)*, *24*(4), 593.

Stats SA. (2007a). *Community survey 2007.* https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/101.

Stats SA. (2007b). *Mortality and causes of death in south africa: Findings from death notification 2007.* http://www.statssa.gov.za/publications/P03093/P030932007.pdf.

Stats SA. (2008). *Mortality and causes of death in south africa: Findings from death notification 2008.* http://www.statssa.gov.za/publications/P03093/P030932007.pdf.

Stats SA. (2009). *Mortality and causes of death in south africa: Findings from death notification 2009.* http://www.statssa.gov.za/publications/P03093/P030932007.pdf.

Stats SA. (2010). *Mortality and causes of death in south africa: Findings from death notification 2010.* http://www.statssa.gov.za/publications/P03093/P030932007.pdf.

Stats SA. (2011a). *Census 2011.* https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/485.

Stats SA. (2011b). *Census 2011 municipal fact sheet.* http://www.statssa.gov.za/census/census_2011/census_products/Census_2011_Municipal_fact_sheet.pdf.

Stats SA. (2011c). *Mortality and causes of death in south africa: Findings from death notification 2011.* http://www.statssa.gov.za/publications/P03093/P030932007.pdf.

Stats SA. (2011d). *The south african mpi.* http://www.statssa.gov.za/publications/Report-03-10-08/Report-03-10-082014.pdf.

Stats SA. (2012). *Mortality and causes of death in south africa: Findings from death notification 2012.* http://www.statssa.gov.za/publications/P03093/P030932007.pdf.

Stats SA. (2013). *Mortality and causes of death in south africa: Findings from death notification 2013.* http://www.statssa.gov.za/publications/P03093/P030932007.pdf.

Stats SA. (2014). *Mortality and causes of death in south africa: Findings from death notification*

2014. http://www.statssa.gov.za/publications/P03093/P030932007.pdf.

Stats SA. (2015a). *Mid year population estimates 1996 - 2015*. http://www.statssa.gov.za/?page_id=1866&PPN=P0302&SCH=7362.

Stats SA. (2015b). *Mortality and causes of death in south africa: Findings from death notification 2015*. http://www.statssa.gov.za/publications/P03093/P030932007.pdf.

Stats SA. (2016a). *Community survey*. https://www.datafirst.uct.ac.za/dataportal/index.php/catalog/611.

Stats SA. (2016b). *Cs 2016 statistical release*. http://cs2016.statssa.gov.za/?page_id=142.

Stats SA. (2016c). *Provincial profile 2016 eastern cape*. http://cs2016.statssa.gov.za/?page_id=142.

Stats SA. (2016d). *Provincial profile 2016 free state*. http://cs2016.statssa.gov.za/?page_id=142.

Stats SA. (2016e). *Provincial profile 2016 gauteng*. http://cs2016.statssa.gov.za/?page_id=142.

Stats SA. (2016f). *Provincial profile 2016 kwazulu-natal*. http://cs2016.statssa.gov.za/?page_id=142.

Stats SA. (2016g). *Provincial profile 2016 limpopo*. http://cs2016.statssa.gov.za/?page_id=142.

Stats SA. (2016h). *Provincial profile 2016 mpumalanga*. http://cs2016.statssa.gov.za/?page_id=142.

Stats SA. (2016i). *Provincial profile 2016 northern cape*. http://cs2016.statssa.gov.za/?page_id=142.

Stats SA. (2016j). *Provincial profile 2016 north west*. http://cs2016.statssa.gov.za/?page_id=142.

Stats SA. (2016k). *Provincial profile 2016 western cape*. http://cs2016.statssa.gov.za/?page_id=142.

System, D. H. I. (2007-2016). *Dhis*.

Tanser, F., Bärnighausen, T., Grapsa, E., Zaidi, J., & Newell, M.-L. (2013). High coverage of art associated with decline in risk of hiv acquisition in rural kwazulu-natal, south africa. *Science*, *339*(6122), 966–971.

UNAIDS. (2020). *South africa*. https://www.unaids.org/en/regionscountries/countries/southafrica.

United Nations Development Programme. (2019). *Human development index*. http://hdr.undp.org/en/indicators/69206#.

Wouters, E., Van Loon, F., Van Rensburg, D., & Meulemans, H. (2009). State of the art: clinical efficacy and improved quality of life in the public antiretroviral therapy program, free state province, south africa. *AIDS care*, *21*(11), 1401–1411.

8 Appendix

Table 10: Data availability and specificity

Variable	District	Province
Mortality rate	AGE	AGE
HIV-related mortality rate	NOT AVAILABLE	AGE, GENDER (15-49)
ARV usage rate	NOT SPECIFIC	AGE
HIV prevalence	NOT SPECIFIC	AGE, GENDER (15-49)
Poverty Headcount	NOT SPECIFIC	NOT SPECIFIC

AGE = age-specific, GENDER = gender-specific

Age-category of gender specificity in parentheses

Table 11: District FE regressions including time dummies

	(1)	(2)	(3)
	Mortality rate	Mortality rate	Mortality rate
	< 15 years old	15-49 years old	50+ years old
ARV rate	-0.0938**	-0.241***	0.243
	(0.0394)	(0.0791)	(0.267)
Lag HIV rate	0.00456**	0.0112**	0.0275*
	(0.00199)	(0.00520)	(0.0150)
Poverty headcount	0.000406	0.0119**	0.00637
	(0.00329)	(0.00590)	(0.0249)
2009	-0.000666***	-0.000207	0.00112
	(0.000130)	(0.000340)	(0.00105)
2010	-0.000484**	6.71e-05	0.000824
	(0.000215)	(0.000520)	(0.00151)
2011	-0.00107***	-0.000675	-0.00104
	(0.000275)	(0.000552)	(0.00160)
2012	-0.00102***	-0.00113*	-0.00210
	(0.000279)	(0.000570)	(0.00161)
2013	-0.000975***	-0.00109	-0.00172
	(0.000315)	(0.000694)	(0.00191)
2014	-0.00107***	-0.00154**	-0.00155
	(0.000313)	(0.000734)	(0.00202)
2015	-0.00115***	-0.00132	0.00139
	(0.000327)	(0.000786)	(0.00219)
2016	-0.00142***	-0.00151*	0.00252
	(0.000342)	(0.000806)	(0.00229)
Constant	0.00409***	0.00762***	0.0262***
	(0.000568)	(0.00146)	(0.00461)
Observations	463	463	463
R-squared	0.616	0.618	0.087
Number of groups	52	52	52

Robust standard errors in parentheses

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

Table 12: Provincial FE regressions per age category including time dummies

	(1)	(2)	(3)
	Mortality rate	Mortality rate	Mortality rate
	< 15 years old	15-49 years old	50+ years old
ARV rate	-0.318***	-0.0734***	-0.0195***
	(0.0289)	(0.00583)	(0.00401)
Lag HIV rate	0.0579***	0.0542***	0.0637***
	(0.0129)	(0.0114)	(0.0142)
Poverty headcount	0.00349	0.00675	-0.00825**
	(0.00362)	(0.00508)	(0.00306)
1999	0.000230***	-4.70e-05	0.000106
	(3.77e-05)	(0.000194)	(5.72e-05)
2000	0.000441***	2.36e-06	0.000251*
	(7.83e-05)	(0.000386)	(0.000124)
2001	0.000589***	2.34e-05	0.000394*
	(0.000121)	(0.000563)	(0.000194)
2002	0.000705***	0.000320	0.000568*
	(0.000140)	(0.000651)	(0.000284)
2003	0.000740***	0.000658	0.000799*
	(0.000173)	(0.000723)	(0.000383)
2004	0.000790***	0.00113	0.00115**
	(0.000202)	(0.000785)	(0.000490)
2005	0.000859**	0.00171*	0.00107*
	(0.000260)	(0.000861)	(0.000564)
2006	0.000793**	0.00180*	0.00113
	(0.000267)	(0.000877)	(0.000624)
2007	0.000620**	0.00103	0.000803
	(0.000236)	(0.000856)	(0.000608)
2008	0.000330	1.98e-05	0.000622
	(0.000229)	(0.000777)	(0.000554)
2009	2.91e-05	-0.000714	0.000131
	(0.000202)	(0.000696)	(0.000517)
2010	0.000168	-0.00101	0.000118
	(0.000244)	(0.000736)	(0.000547)
2011	0.000428	-0.00115	4.56e-05
	(0.000254)	(0.000756)	(0.000566)
2012	0.000632**	-0.00124	2.84e-05
	(0.000263)	(0.000783)	(0.000568)
2013	0.000917**	-0.00115	-9.47e-05
	(0.000276)	(0.000813)	(0.000590)
2014	0.000940***	-0.000827	-0.000449
	(0.000264)	(0.000860)	(0.000611)
2015	0.000980***	-0.000596	-2.69e-05
	(0.000260)	(0.000900)	(0.000666)
Constant	0.000930	-0.00274	0.00124*
	(0.000675)	(0.00163)	(0.000647)
Observations	162	162	162
R-squared	0.950	0.945	0.866
Number of groups	9	9	9

Robust standard errors in parentheses

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

Table 13: Provincial FE regressions per gender including time dummies

	(1)	(2)
	Mortality Rate	Mortality Rate
	Men 15-49 years old	Women 15-49 years old
ARV rate	-0.0606***	-0.0846***
	(0.00515)	(0.00794)
1999	2.89e-05	-0.000135
	(0.000154)	(0.000219)
2000	0.000141	-0.000155
	(0.000306)	(0.000437)
2001	0.000238	-0.000211
	(0.000443)	(0.000640)
2002	0.000458	0.000128
	(0.000519)	(0.000747)
2003	0.000690	0.000541
	(0.000590)	(0.000835)
2004	0.00100	0.00114
	(0.000655)	(0.000909)
2005	0.00151*	0.00180
	(0.000722)	(0.000991)
2006	0.00162*	0.00188
	(0.000743)	(0.00101)
2007	0.00110	0.000860
	(0.000726)	(0.000974)
2008	0.000661	-0.000594
	(0.000697)	(0.000862)
2009	1.45e-05	-0.00143
	(0.000606)	(0.000810)
2010	-0.000243	-0.00178*
	(0.000641)	(0.000846)
2011	-0.000384	-0.00194*
	(0.000598)	(0.000895)
2012	-0.000519	-0.00204*
	(0.000581)	(0.000961)
2013	-0.000487	-0.00193*
	(0.000588)	(0.00102)
2014	-0.000148	-0.00163
	(0.000645)	(0.00105)
2015	5.15e-05	-0.00139
	(0.000687)	(0.00109)
Lag HIV rate	0.0603***	0.0517***
	(0.0116)	(0.0111)
Poverty headcount	0.00369	0.0101
	(0.00439)	(0.00567)
Constant	-0.00194	-0.00375*
	(0.00127)	(0.00189)
Observations	162	162
R-squared	0.941	0.944
Number of code	9	9

Robust standard errors in parentheses

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

Table 14: Provincial robustness regressions using HIV prevalence per gender

	(1)	(2)
	Mortality Rate	Mortality Rate
	Men 15-49 years old	Women 15-49 years old
ARV rate	-0.0558***	-0.0779***
	(0.00496)	(0.00764)
HIV prevalence	0.0534***	0.0471***
	(0.0146)	(0.0107)
Poverty headcount	0.000934	0.00589
	(0.00490)	(0.00509)
Constant	-0.00169	-0.00332*
	(0.00148)	(0.00168)
Observations	171	171
R-squared	0.929	0.938
Number of groups	9	9

Robust standard errors in parentheses

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

Table 15: Provincial robustness regression overall mortality per gender

	(1)	(2)
	Mortality Rate	Mortality Rate
	Men 15-49 years old	Women 15-49 years old
ARV rate	-0.0662***	-0.0804***
	(0.00791)	(0.0113)
Lag HIV prevalence	0.0330	0.0476*
	(0.0280)	(0.0232)
Poverty headcount	-0.00320	0.0144
	(0.00809)	(0.0114)
Constant	0.00595*	-0.00178
	(0.00265)	(0.00372)
Observations	162	162
R-squared	0.906	0.913
Number of code	9	9

Robust standard errors in parentheses

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