# The effect of proximity to the northern Namibian borders on smoking prevalence among Namibian men

**Tuyenikeumbo Rose Peter**

**KVNTUY001**



**Supervised by**

**Prof. Corne van Walbeek**

Presented in partial fulfilment of the requirements of the Masters of Commerce in Applied Economics Degree at the University of Cape Town, 15 February 2016

## Acknowledgements

I would like to thank my family, Mom, Dad, Kauly, Peter, Loide and Josh. Their pep talks at awkward hours of the day, random messages of encouragement, prayers, love, generosity, patience, and proof reading have kept me going. I would not be here without them.

Thank you to my supervisor, Corné van Walbeek, for guiding, supporting and motivating me through two very difficult academic years. His understanding and kindness helped me not to lose confidence in myself.

Many thanks to Evan Blecher and Hana Ross for their assistance in conceptualizing this research topic. Thank you to Tim Brophy for his patience in assisting me with the mapping software, which made a significant portion of this thesis possible.

I would also like to thank my best friend, Xungi, for taking the time out of her work schedule to proof read and edit this entire document; I truly appreciate it.

Thank you to the Bill & Melinda Gates Foundation and the National Research Foundation for their funding of my research.

Finally, I would like to thank God for blessing me with this opportunity and bringing all the people into my life who made this possible. I am truly grateful.

# Plagiarism declaration

1. I know that plagiarism is wrong. Plagiarism is to use another's work and pretend that it is one's own.
2. I have used the Harvard convention for citation and referencing. Each contribution to, and quotation in this thesis from the work(s) of other people has been attributed, and has been cited and referenced.
3. This thesis is my own work.
4. I have not allowed, and will not allow, anyone to copy my work with the intention of passing it off as his or her own work.
5. I acknowledge that copying someone work or part of it, is wrong, and declare that this is my own work.

Signed by candidate

Signature Removed

Signature:

Date: 15 February 2016

# Abstract

This thesis draws on the findings of research in the North America and EU, which shows that price differences of cigarettes between neighboring countries or states (in the Unites States of America) are associated with higher odds of cross border purchase and may lead to smuggling of cigarettes. This results in the tax revenue generation and public health aims of tobacco control policy, through tobacco taxation, being undermined. The Namibian Demographic Health Survey (NDHS) 2013 data is used to assess whether the probability of smoking among Wambo, Lozi and Kwangali Namibian men, living within 150km of the Angolan or Zambian borders, is affected by their proximity to these borders, given that cigarettes are cheaper in Angola and Zambia, than in Namibia. Logistic regressions are used to assess whether proximity to these borders has an effect on the likelihood of smoking, and smoking intensity. The results show that proximity to the border has no statistically significant effect on the probability of smoking or intensity among this group. This may mean that the Namibian government can in fact can pursue more aggressive tobacco taxes, to reduce consumption of tobacco products, without encouraging illicit trade, cross border purchases, or compromising its public health agenda.

# Table of Contents

# 1. Introduction

Tobacco use is the leading cause of preventable disease and death in the world (International Tobacco Control (ITC) Project, 2014b). Of the six million people killed by tobacco related illness around the world, 80% of those people live in low and middle income countries (ITC Project, 2014b). Trends in tobacco use over the past few decades have shown that the prevalence of tobacco use has been declining in developed nations, whereas in low and middle income countries it is increasing due to the marketing efforts of tobacco companies (Guindon & Boisclair, 2003; Pampel, 2008; World Health Organisation, 2011). This is especially true in Africa, which has a relatively low smoking prevalence of only 12% compared to other regions, but has been seeing an increase in smoking prevalence in recent decades (Guindon & Boisclair, 2003; John, Mamudu, & Liber, 2012; ITC Project, 2014b). Besides the global reach and influence of the tobacco industry, this increase in smoking prevalence is said to be caused by increased urbanisation, growing incomes and low levels of literacy (John, et al., 2012). Due to the paucity of high quality, nationally representative data on tobacco use in Africa, it is difficult to obtain reliable estimates of the prevalence of tobacco smoking in many African countries. However Bilano et al. (2015) have estimated that, given the trends from 2000 to 2010 in 40 African countries considered in their study, smoking prevalence among men in Africa is predicted to increase in 15 of these countries from 2010 to 2025. This is considering that the trends already show that the prevalence of smoking has increased between 2000 and 2010 in 25 out of the 40 countries under consideration (Bilano, et al., 2015). Considering that tobacco use is the leading cause of preventable death globally and that this burden is disproportionately borne by low and middle income countries (World Health Organisation, 2011), Bilano et al's (2015) findings do not bode well for Africa.

The World Health Organisation Framework Convention on Tobacco Control (WHO FCTC) is the first global public health treaty and was developed with the intention of stemming the "globalisation of the tobacco epidemic" (World Health Organisation, 2015b). The WHO FCTC aims to do this by addressing the causes of the tobacco epidemic such as "cross border effects", which include "trade liberalization and direct foreign investment, tobacco advertising, promotion and sponsorship beyond national borders, and illicit trade in tobacco products" (World Health Organisation, 2015b). Namibian, Angola and Zambia are all signatories to this treaty (United Nations, 2016).

Namibia is still considered to be in the early stages of the tobacco epidemic, as it has a cigarette or pipe smoking prevalence rate of only 19% for men and 5% for women (Tam & van Walbeek, 2013; The Nambia Ministry of Health and Social Services (MoHSS) and ICF International, 2014). When Namibia became an independent country in 1990, there was no existing tobacco control legislation to speak of (Tam & van Walbeek, 2013). Discussions around the implementation of tobacco control legislation started in 1991, but due to threats of legal action from the tobacco industry and a lack of capacity within the government, it was only in 2010 that Namibia's first tobacco control act came into place (Tam & van Walbeek, 2013). Although the Tobacco Products Control Act of 2010 was modelled around South Africa's tobacco control legislation, much of the Act was informed by the WHO FCTC, such as sections related to youth smoking, passive smoking and illicit trade (Tam & van Walbeek, 2013). This legislation bans smoking in public places, and advertising of tobacco products on national television, radio and print media (Tumwine, 2011). As of 1 April 2015 pictorial warnings have also been introduced on the packaging of tobacco products (World Health Organisation Regional Office for Africa, 2015).

Angola has an adult smoking prevalence rate of 16.7% among men and 1.6% among women. Although the Angolan government is a signatory to the WHO FCTC, has specific national objectives for tobacco control, and a national agency dedicated to tobacco control, their tobacco control policy is limited. There are no legal restrictions in place to protect people from smoking in public places or adequate health warnings on packaging. (Tumwine, 2011; The Tobacco Atlas, 2015; World Health Organisation, 2015c).

The adult smoking prevalence in Zambia is estimated to be 23.8% among men and 0.4% among women (World Health Organisation, 2015d). In 1992 the Zambian government implemented the Public Health (Tobacco) Regulations which banned smoking in public places, introduced English text warnings on tobacco products, banned the selling of tobacco products to those under 16 years of age, and banned direct and indirect commercial advertising, except for direct advertising to the public (ITC Project, 2015). Since Zambia's ratification of the WHO FCTC, work has been done to strengthen the 1992 Public Health Regulations (ITC Project, 2014b). Zambia also experienced push back from tobacco companies in drafting more comprehensive tobacco control legislation (Moonga, 2011). Although it went against the recommendations of the WHO FCTC, tobacco companies were consulted and involved in the drafting of the Tobacco Products Control Bill of 2010 (Moonga, 2011). Zambia is also a tobacco producing country and many rural farmers make their livelihoods farming tobacco (ITC Project, 2014b).

For fear of worsening poverty and losing the foreign exchange earned through the exporting of tobacco, the Zambian government has been hesitant to ban the production of tobacco (News Ghana, 2016). Although there are programs in place to support Zambian tobacco farmers in switching to farming other crops, and many farmers express a desire to make this switch, few know of these support programs (ITC Project, 2014b).

Table 1 shows the prices of different categories of cigarettes in US Dollars (US$) for Namibia, Angola and Zambia. Except for the premium brand of cigarettes, cigarettes in Namibia are significantly more expensive than both Angola and Zambia. For the premium brand, Zambia has the highest cigarette price at US$4.89 at the official exchange rate, compared to US$3.55 and US$2.06 in Namibia and Angola respectively. The cheapest brand of cigarettes in Namibia costs just over one and a half times as much as those in Zambia and only US$0.05 more than the Angolan equivalent at the prevailing US$ official exchange rate. The most popular brand of cigarettes in Namibia is also more than two and a half times more expensive than the most popular Zambian brand, and almost twice as expensive as Angola's most popular brand at the official US$ exchange rate. This significant price differential can be partially explained by the total tax rate on the most popular brand of cigarettes in table 2. As one can see, the total tax on cigarettes is 9.1% and 11.4% higher in Namibia than it is in Angola and Zambia, respectively.

**Table 1: Cigarette prices in US Dollars**

| Country | Price of 20-cigarette pack of Marlboro or other premium brand | | Price of a 20-cigarette pack of the cheapest brand | | Price of 20-cigarette pack of the most sold brand | |
|---|---|---|---|---|---|---|
| | International US$ (at purchasing power parity[1]) | US$ at the official exchange rate | International US$ (at purchasing power parity) | US$ at the official exchange rate | International US$ (at purchasing power parity) | US$ at the official exchange rate |
| Namibia | 5.46 | 3.55 | 2.44 | 1.59 | 5.74 | 3.74 |
| Angola | 2.18 | 2.06 | 1.63 | 1.54 | 2.18 | 2.06 |
| Zambia | 6.00 | 4.89 | 1.20 | 0.98 | 1.80 | 1.47 |

Source: World Health Organisation, 2015a

[1] Purchasing power parity refers to the adjustment required in the exchange rate of two currencies to make their purchasing power comparable.

**Table 2: Taxes as a percentage of the price of the most sold brand**

| Country | Specific Tax | Ad Valorem Tax | Value Added Tax (VAT) | Other Taxes | Total Tax |
|---|---|---|---|---|---|
| Namibia | 29.0% | 0.0% | 3.8% | 0.0% | 32.8% |
| Angola | 0.0% | 0.0% | 22.9% | 0.8% | 23.7% |
| Zambia | 0.0% | 20.0% | 1.4% | 0.0% | 21.4% |

Source: World Health Organisation, 2015a

The importing of cigarettes and other tobacco products into Namibia is governed by the Customs and Excise Act of 1998. Its provisions pertaining to cigarettes are mainly related to the prohibition and manufacturing of cigarettes over a certain weight[2] in Namibia (Customs and Excise Act 20 , 1998). There are also limitations on the quantities of tobacco products which visitors to Namibia, returning citizens and other passengers to Namibia are allowed to bring into the country through duty free allowances. Citizens of Southern African Customs Union (SACU) member states, namely Namibia, South Africa, Botswana, Lesotho and Swaziland (SACU, 2013), are allowed to bring in higher volumes of tobacco products compared to those who are from non SACU countries. SACU citizens are allowed 400 cigarettes, 50 cigars, and 250 grams of cigarette or piped tobacco, whereas non SACU members are allowed 200 cigarettes and 20 cigars, but the same amount of smoking and pipe tobacco (Namibia Trade Information Portal, 2015). These duty free allowances make it possible for travellers to buy cheaper cigarettes and other tobacco products in neighbouring countries and legally bring them into the country for personal consumption. Although this is not illegal, it is a way in which consumers can buy tobacco products without paying the full tax for them.

Namibia's northern border is of primary interest as it is the most densely populated area of the country (Namibia Statistics Agency (NSA), 2014) and the porosity of these borders is well documented. Since Angola and Zambia are the only countries which border Namibia in the north, only these countries were chosen for the analysis. Namibia's porous northern borders make it relatively easy to bring goods into and out of the country without detection. Namibia's borders were drawn up at the Berlin Conference in 1884, when Namibia was a German colony (Griffiths, 1986; Zeller, 2010). Like many African countries, Namibia's borders are strange in that, besides those demarcated by bodies of water or other geographic features, many of them

[2] The Customs and Excise Act of 1998 prohibits the importing and manufacture of cigarettes over 2kg per 1000 cigarettes (Customs and Excise Act 20 , 1998).

are straight lines (Griffiths, 1986). These borders were drawn up without any consideration of the communities who lived in Namibia or its neighbouring countries (Griffiths, 1986; Zeller, 2010). As a result, Namibia's northern border cuts through many communities and ethnic groups (Griffiths, 1986; Zeller, 2010). It is for this reason that this northern border is very porous – people easily move across it to graze cattle, trade and visit their families. Except for the inclusion of Walvis Bay into Namibia (formerly it was a part of the Republic of South Africa), these borders were not changed after Namibia's independence in 1990 (Simon, 1998). Although there are official border crossings in the Ohangwena and Omusati regions of Namibia, the border can be crossed at any point (Gueye, et al., 2014). Namibian or Angolan citizens who live in settlements close to the border can be granted border residency cards which allow them to be within 60km of either country without a passport (Namibian Sun, 2013; Gueye, et al., 2014). Although there is no formal or legal mechanism for Namibian and Zambian citizens to cross the Zambia-Namibia border as easily, this border is also very porous allowing ease of movement without formal detection between the two countries (New Era, 2007; Zeller, 2010).

The porousness of the Namibian borders with Zambia and Angola, combined with the significant price difference between Namibia and its northern neighbours would make cross border purchases or smuggling of cheaper, lower tax cigarettes very easy. According to the literature, these conditions create an incentive for the purchase of cheap, potentially illegal cigarettes. This thesis aims to investigate whether the effects of easy access to cheap cigarettes from Angola and Zambia affects the likelihood of smoking among men who live sufficiently close to the northern borders of Namibia, using data from the Namibian Demographic Health Survey (NDHS) 2013. Only men are included in this study, as there are too few female smokers in the NDHS dataset, but this will be explained in further detail in section 3. It is hypothesised that proximity to the border should have an effect on the probability of smoking and this effect should be positive – meaning that being close to the border should increase one's likelihood of smoking.

The rest of this paper is organised as follows: section two reviews and discusses the relevant literature, and section three presents the methodology used to analyse the NDHS data and some descriptive statistics. Results are presented and discussed in section four. Section five discusses these results, their limitations and policy implications, and section six concludes.

# 2. Literature review

It has been argued that the real price of cigarettes leads to a reduction in the prevalence of smoking, however the opposite is also true. Consumers treat cheaper, low-tax cigarettes as substitutes for more expensive, full-tax cigarettes. If cigarette prices are increased, but one has access to cheaper cigarettes elsewhere, then smokers will continue to buy cigarettes where they are cheaper. Consumers make this decision by looking at the "full price" of cigarettes, which is not only made up of monetary price which the buyer pays the seller for the cigarettes, but also the "transaction cost" (Ross, 2015). The transaction cost consists of the cost of the convenience of obtaining cigarettes, in terms of time and travel distance required to access the product, and the risk which comes with buying and consuming the cheaper and potentially illegal cigarettes (Ross, 2015). If the total cost of a cross border purchased or illegally smuggled cigarettes is lower than the cost of full-tax, legal cigarettes, this may result in the undermining of tobacco control policies. This review of literature aims to look at what the literature says about the relationship between cigarette prices and demand. It then looks at how this relationship is affected by access to cheaper cigarettes through cross border purchase of cigarettes, both legal and illegal. The literature, or lack thereof, on cross border smuggling in Africa is then discussed, after which different methodologies of measuring cross border purchase or smuggling, for the purpose of determining the size of tax avoidance or evasion, are explored.

There are many ways in which people attempt to avoid paying higher cigarettes prices, but they are essentially ways of avoiding or evading the taxes which make cigarettes more expensive. Tax avoidance is not illegal. It can take the form of buying cheaper cigarettes in line with tax and customs regulations, for instance buying from duty free vendors and cross-border shopping, among other methods. These methods involve paying some taxes, but not all. Tax evasion, on the other hand, is illegal and smuggling cigarettes is a form of tax evasion. Small scale smuggling is also called "bootlegging" and involves moving cigarettes across borders in quantities which are in excess of allowable limits. These products are often smuggled with the intention of being sold in a different jurisdiction without paying the required amount of tax, at a lower price. Large scale smuggling operations are usually run by criminal networks. They generally involve very large quantities of cigarettes and result in complete tax avoidance. (Ross, 2015).

The relationship between price and cigarette consumption has been studied extensively in tobacco economics literature. Increasing the real price of cigarettes leads to a reduction in consumption, despite the addictive nature of cigarettes (Chaloupka & Warner, 2000; Chaloupka, Yurekli, & Fong, 2012). Increasing the real (inflation adjusted) price of cigarettes by 10% leads to a reduction of consumption of cigarettes by 4-8% (Blecher & van Walbeek, 2004), however this effect is lower in high income countries compared to low income ones (Saloojee, 1995; Chaloupka & Warner, 2000; Blecher & van Walbeek, 2004; Jha, Chaloupka, Corrao, & Jacob, 2006). The effect of real price on cigarettes consumption, or the price elasticity of demand, also differs according to age, socioeconomic status, education, and even ethnicity or race (Townsend, Roderick, & Cooper, 1994; Centres for Diesease Control and Prevention, 1998; Jha, Chaloupka, Corrao, & Jacob, 2006).

In the African context, South Africa is probably the most well studied example of the impact increasing the real price of cigarettes has on consumption. Van Walbeek (2002) found that in South Africa the implementation of anti-smoking legislation, which included greater taxation of tobacco products, led to an increase in the real price of cigarettes by 93% between 1993 and 2000. This price increase led to a reduction in aggregate cigarette consumption of 22%, along with a fall in smoking prevalence in the adult population from 32% to 28% (van Walbeek, 2002; Jha & Peto, 2014). Chelwa et al (2015) took this finding a step further and showed that even though the prevalence of tobacco use was already declining, the substantial increase in the real price of cigarettes caused by the introduction of the anti-smoking tobacco legislation was the main driving force behind the reduction in smoking prevalence. They found that per capita cigarette consumption in South Africa was 36% lower than it would have been, had the tax increases not been implemented (Chelwa, et al., 2015).

The effect of price differentials, between the regions of countries and countries themselves, on the demand for cigarettes is well documented in the literature, particularly in North America and the European Union (EU). These studies mainly look at the effects on demand due to price differences which occur as a result of differences in taxes levied on cigarettes between neighbouring states within the United States (US), or EU member states. Many of the US studies try to tease out how much of the change in price elasticity of demand can be attributed to “smuggling” of cigarettes from states with lower taxes, as opposed to the actual change in price of cigarettes in the home state. The idea is that if cigarettes are cheaper in state A than in state B due to lower taxes on cigarettes, consumers from state B who live sufficiently close to the border of state A will buy cigarettes from state A instead. As a result, when price elasticities of demand are calculated they may be biased, as they do not take into account the effect of cigarette smuggling from neighbouring states. This "casual smuggling" can lead to loss of tax revenue in states with higher taxes on cigarettes, and discourage consumers from smoking less, undermining both the revenue generation and the public health intervention purposes of tobacco taxation. (Baltagi & Levin, 1986; Baltagi & Goel, 1987; Lovenheim, 2008; Ross, 2015).

One of the most documented cases of the effect of cigarette taxation on smuggling was between the US and Canada. In the late 1980s and early 1990s, Canada increased taxes on cigarettes which lead to large price differences between cigarettes sold in Canada and the US. The relatively low price of US cigarettes, coupled with weak border controls between the US and Canada, allowed for an increase in smuggling of cigarettes from the US into Canada in the early 1990s. (Sweanor & Martial, 1994; Chaloupka & Warner, 2000)

Agaku et al. (2015) have found that in spite of the EU's highly successful tax harmonisation structures, there is still a great deal of variation in cigarette prices among member states due to remaining differences in the taxation on tobacco products. These differences in prices result in consumers engaging in "tax avoidance strategies", such as purchasing cigarettes in countries where they are cheaper. The authors found that, according to Eurobarometer 385 survey, as many as 26.2% of current cigarette smokers in the EU purchased tobacco in another country in the last twelve months. Of those who engaged in these cross border purchases, 56.3% of respondents said they did it because cigarettes were cheaper in the other country. They also found that EU countries that were bordered by at least one country, where the price for the most popular price category (MPPC) of cigarette was more than 25% lower, had the highest prevalence of cross-border tobacco purchase compared to those that were not. In addition to this, they found a positive relationship between the odds of cross border purchase, and the size of the price difference in the MPPC of cigarettes in the home country and neighbouring country. In other words the larger the price difference in MPPC between two countries that neighbour each other, the higher the odds of cross border purchases happening. One of the main reasons taxes, such as value added tax and excise tax, are applied to tobacco products is to reduce consumption and prevalence of tobacco use by increasing the real price of cigarettes. As a result, substantial price differentials of cigarettes between EU member states may mute the effectiveness of tobacco control legislation. (Agaku, et al., 2015).

Since Agaku et al (2015) find that differences in prices of tobacco products between neighbouring countries is associated with higher odds of cross border purchases, one may be tempted to argue in favour of tax harmonisation for countries which find themselves in this position. Although this relationship may exist, there is no proof that greater harmonisation of taxes on tobacco products reduces illicit trade in tobacco products. However, harmonisation of excise taxes in SACU has had a positive effect in reinforcing public health agendas. The excise tax on tobacco products is set annually by South Africa at a rate of 52% of the recommended retail price, which includes the uniform specific excise tax plus VAT. This has worked for SACU member states, because the dominant country in the customs union, South Africa, pursues an aggressive tobacco tax policy as part of its tobacco control agenda. By extension, SACU member states have benefited from this. The opposite would be true if South Africa were not pursuing this policy, because other SACU member states would not have had the freedom to use taxes to fulfil their public health agendas, because of harmonised excise taxes. (Blecher & Drope, 2014).

Not much has been written about cross border purchases or smuggling in African countries. Although there have been many reports of illicit trade in tobacco in the media (Umraw, 2015; Koyana, 2015; New Era, 2015; Lusaka Times, 2015; Lusaka Voice, 2015; Gallet, 2015; Ezeamalu, 2015; Gbubemi, 2015; Sasman, 2013), there is a shortage of academic studies calculating the actual impact of cigarette smuggling among African countries. Blecher (2010) attempts to estimate the size of the illegal cigarette market in South Africa using data from van Walbeek (2005), the All Media Product Survey (AMPS) and population data from Statistics South Africa. Blecher (2010) shows that the size of the illicit cigarette market in South Africa is exaggerated by tobacco companies, in an attempt to undermine the use of tobacco control policies which advocate for increased taxation of tobacco products. Tobacco firms often use the increase in cross border purchase of cigarettes to argue against the increase in taxes on tobacco products. They argue that higher taxes on tobacco products, which result in higher prices, encourage cross border purchases and illicit trade, which results in losses of tax revenue, therefor higher taxes on tobacco products are counterproductive (Blecher, 2010). However, public health experts argue that although high income countries have higher cigarette prices, they have lower levels of cigarette smuggling than lower income countries (Blecher, 2010). In fact, even in Europe, northern European countries with higher cigarette prices still experience lower levels of smuggling than their southern European counterparts who have lower cigarette prices (Blecher, 2010).

Hanna Ross's (2015) methodological guide to "Understanding and measuring cigarette tax avoidance and evasion" and David Merriman's (2012) paper, "Understand, Measure and Combat Tobacco Smuggling" provide useful discussions of different methods which have been developed to track cross border purchases of and illegal trade in tobacco products in order to measure tax avoidance and tax evasion. These methods include designing and/or conducting surveys of tobacco users in order to inspect cigarette packs and to find out the source of their cigarette packets, and observational studies in the form of collection of cigarette packs for examination (Merriman, 2012; Ross, 2015). "Gap-analysis" is also a useful method of measuring tax avoidance or evasion (Ross, 2015). Gap analysis measures the "difference between estimated consumption of cigarettes at national and/or local levels and tax-paid sales for the corresponding area" (Ross, 2015). Econometric modelling can also be used to estimate demand functions using regression analysis, with micro or macro level data in an attempt to infer tax avoidance/evasions (Merriman, 2012; Stoklosa & Ross, 2013; Ross, 2015). Official tax paid sales data is used to measure demand, and it is estimated as a function of variables which are known to affect demand (Ross, 2015). This method is good for measuring how sensitive tax avoidance/evasion is to changes in taxes or other factors which may influence them (Ross, 2015). It can also tease out how much of changes in sales can be attributed to tax avoidance versus evasion, if the demand function is specified correctly (Ross, 2015). Key informant interviews and the monitoring of trade are also other ways of measuring tax evasion/avoidance through illicit trade or legal cross border purchases (Merriman, 2012; Ross, 2015).

Given the literature reviewed here, one can see that nothing has been written about the effects of cross border smuggling on the populations living near neighbouring borders of countries with cheaper cigarettes in Africa. This is most likely due to a paucity of data. It is often very difficult to obtain reliable estimates of the presence and the scope of tax avoidance or evasion in cigarette consumption (Ross, 2015). This is usually due to lack of reliable data on who buys low tax cigarettes, or where there is data, it may not always be publicly available (Ross, 2015). It is for this reason that researchers often need to use innovative methods on existing datasets which were compiled for other purposes to try and measure the effects of access to cheap cigarettes (Ross, 2015). This is what has been done in this paper.

Given the findings of Agaku et al (2015), there is reason to believe that the substantial difference between the MPPC between Namibia and, Angola and Zambia as presented in table 1, implies higher odds of cross border purchase of cigarettes or smuggling. Creative methods have had to be employed on the Namibian Demographic Health Survey (NDHS) 2013 to try and see whether the effect of proximity to countries with cheap cigarettes can be measured. These methods are described in greater detail in section 3.

# 3. Methodology

## 3.1 Data

The data source of this study is the Namibian Demographic Health Survey (NDHS) 2013. The NDHS 2013 is part of the global Demographic Health Survey (DHS) program sponsored by the United States Agency for International Development (USAID), with the aim of "providing demographic, socioeconomic, and health data necessary for policymaking, planning, monitoring, and evaluation of national health and population programmes". The NDHS 2013 is Namibia's fourth comprehensive, nationally representative, demographic health survey. (MoHSS and ICF International, 2014).

The NDHS has a sample of 14499 individuals, in 9845 households, sampled from all 13 administrative regions of Namibia. Three different questionnaires were used for data collection, the Household Questionnaire, the Men's Questionnaire and the Women's Questionnaire. The Men's Questionnaire was only administered to men aged 15 to 64 in half of the households, so the sample of men in the dataset is much lower than that of the women. The Women's Questionnaire was administered to women aged 15 to 49 in all selected household, but only half of the households were selected to administer the questionnaire to women aged 50 to 64. (MoHSS and ICF International, 2014).

In the DHS questionnaire there are five questions about the use of tobacco products. These questions relate to whether one smokes cigarettes and how many have been smoked in the last 24 hours, the types of tobacco products used (such as, snuff, chewing tobacco, pipe tobacco and others) and whether these products are used daily or not, and lastly the age at which one started using tobacco products daily (MoHSS and ICF International, 2014). Only data on whether respondents are smokers or not, and the number of cigarettes smoked in the last 24 hours is used in this study. Other variables of interest in the NDHS are those associated with tobacco use such as age, gender, education, socioeconomic status, ethnicity, marital status, religion, place of residence and GPS coordinates of each cluster.

One of the main limitations of the NDHS's tobacco use data is that there is a very small sample of female smokers. Only 4% of women aged 15-49 (470 individuals) smoke (MoHSS and ICF International, 2014).

Considering that only households which are sufficiently close to the Angolan and Zambian borders are of interest, the sample is restricted to households within 150km of the Angolan or Zambian borders. The sample was restricted further to include only Wambo, Kwangali and Lozi men. Other language groups were excluded, because too few were sampled to be able to make valid inferences from them by ethnic group, and women were excluded due to the bias which may be caused by the small number of female smokers in the sample. See appendix B for a breakdown of smokers by language spoken and gender. Of the 3768 women who could be included in the sample (606 Kwangali, 362 Lozi and 2800 Wambo women), because they live within 150km of either the Angolan or Zambian border, only 40 of them are smokers. Therefore, if women were included in the sample, the proportion of smokers would fall from 12.18% (in a men only sample) to 4.48%[3] (in a sample with both men and women). A smoking prevalence rate of only 4.48% in the sample would constitute a rare event. Given that maximum likelihood methods are used in estimating the logistic model, having such a small number of cases for an event occurring in a sample of over 4000 people may bias the results significantly, as logistic regressions can significantly underestimate the probability of rare events (King & Zeng, 2001). According to King and Zeng (2001), even with very large sample sizes the biases in the probabilities can still be quite substantial. As a result, the sample size is reduced from 14499 men and women to only 1676 men.

## 3.2 Mapping

The NDHS data contains information on the GPS coordinates of households at cluster level. GPS coordinates are collected using GPS receivers and are accurate to approximately 15 to 20 metres. Where clusters do not have GPS readings, DHS surveyors obtained coordinates from paper maps, gazetteers of settlement names or pre-existing census data. To protect the identity

[3] A sample with men and women would consist of 5442 individuals, of whom 244 are smokers.

of respondents, the GPS coordinates are randomly displaced so that urban clusters contain between 0 – 2 kilometres (km) of error and rural clusters contain between 0 – 5km of positional error, with a further 1% in rural clusters displaced by a maximum of 10km. For this reason, the GPS coordinates cannot be used to calculate exact distances between points, but they can be used to calculate distances as intervals. It should be noted that these displacements have been restricted so that the GPS coordinates remain within the national borders and the administrative regions of the country for the DHS survey. (The DHS Program, 2013).

For this analysis the variables describing the distances to borders are categorical variables. Distances are categorised as 0-25km (category one), 25-50km (category two), and 50-150km (category three) from the borders. Given the displacement of the GPS coordinates, it may be that the distance from the border which has been calculated in the data, may not represent the actual distance from the border for the observations. It may be that due to the positional error caused by displacement, some observations may be miscategorised. For instance, observations which are in reality only 23km from the Angolan border should fall into category one of the distance variable. However, since GPS coordinated are displaced by 0-5km in rural areas, this observation may be displaced into category two. Assuming that positional error is normally distributed, it is estimated that about 4.6% of households in the sample of men under consideration fall into this category. This means that the interpretation of results based on these distance variables is limited by measurement error. See Appendix C for a detailed explanation of how this estimation has been made.

QGIS mapping software was used to obtain the points along the northern border of Namibia. Polygon shapefiles were downloaded from the DIVA-GIS website for the Namibian, Angolan and Zambian administrative areas (DIVA-GIS, n.d.). A detailed description of the methodology used to generate the distance variables from these polygon shapefiles is provided in Appendix D.

### 3.3 Descriptive statistics

Tables 5, 6 and 7 present descriptive statistics of the variables of interest in the model under consideration. These statistics have been included to aid in understanding the distribution of the variables of interest in the sample. Although women are not included in the sample under analysis, they are included in the descriptive statistics to give greater context to the statistics presented for men.

These results have been weighted to ensure suitable representation and to correct for differences in response rates. This dataset contains two types of weights, household and individual weights. Since this analysis is at the individual level, only the individual weights have been used. Individual weights are calculated as the household weight, multiplied by the inverse of the response rate of a particular individual in the sample. The individual weights are also divided into weights for males and females. As a result, descriptive statistics for men and women will be presented separately. In the NDHS dataset, individual weights are only calculated for female respondents aged between 15 and 49 years old, therefore weighted descriptive statistics can only considered for women who fall into this age category. This is not the case for men, as weights are calculated for all age groups. (Shea & Guillermo, 2006).

The sample weights provided in the NDHS are not appropriate for estimating relationships, such as regression analysis and correlation coefficients (Shea & Guillermo, 2006). The use of the NDHS weights will also bias the estimates of confidence intervals, because the number of weighted cases will be taken to produce the confidence interval and not actual number of observations (Shea & Guillermo, 2006). For these reasons, none of the regressions in this study have been weighted.

Table 5: Table of descriptive statistics

The sample analysed in the following tables consists only of Lozi, Kwangali and Wambo people who live within 150km of the Angolan or Zambian borders, they are referred to as the "restricted" sample. Smokers are those who self-report currently smoking cigarettes (MoHSS and ICF International, 2014; MEASURE DHS/ICF International, 2013). The proportion of women who smoke in this restricted sample, is significantly lower than that of men; 1.05% for women as opposed to 12.18% for men. The prevalence of smoking among women and men in this restricted sample, is much lower than that of the NDHS sample as a whole, which is 4% for women and 18.6% for men (MoHSS and ICF International, 2014). This is because, of the administrative regions included in the restricted sample, about half of them have among the lowest smoking prevalence rates for men and women in the entire country.

About a quarter of women and almost a third of men are aged between 15 and 19 years old. It appears that this restricted sample is fairly young, as just over 57% of women and about 65% of men are under 30 years old. Only 3.75% of men in this analysis are over the age of 50. As there are no individual weights for women over the age of 49, no women of this age category included in this table. There about 100 men who have not been assigned an age category due to non-response.

Anyone who has defined their occupation as anything other than "not working" is defined as employed, for the purposes of this analysis. Only those who have reported that they have worked in the past 12 months have defined occupation codes, therefore this variable only includes these respondents. Besides not working, respondents' occupations are professional/technical/managerial, clerical, sales, self-employed agricultural, agricultural employee, services, skilled manual, unskilled manual and other (MEASURE DHS/ICF International, 2013). Using this definition, the unemployment rate is high, at 64.85% for women and 53.78% for men. The unemployment rate in this sample, is also higher than that of the survey sample as a whole, when the same definition is applied. In the whole NDHS sample 54.3% of women and 38.5% of men were not employed in the 12 months preceding the interview (MoHSS and ICF International, 2014). This is most likely due to the fact that the regions which are included in the sample for this study, have higher unemployment rates than the national average of the NDHS sample. These regions are Zambezi, Kavango, Kunene, Ohangwena, Omusati, Oshana and Oshikoto[4]. All these regions have higher unemployment rates for men than the NDHS sample average for men, and all except the Oshana and Oshikoto regions have higher unemployment rates for women, too (MoHSS and ICF International, 2014).

The wealth index is a "composite" measure of the living standard of a household. The DHS uses principal component analysis to place respondents on a continuous scale of relative wealth. Ownership of assets, such as bicycles and televisions, along with materials used to construct the house, and access to water are used to construct this measure. Household are then divided into quintiles from poorest to wealthiest. The distribution along the wealth index between quintiles is similar for men and women. When the top two quintiles are combined, just over 20% of respondents fall into this group, as opposed to over 50% of respondents when the bottom two quintiles are combined. This is because the regions which fall within the 150km cut off from the northern Namibian borders are predominantly rural, and therefore they have a higher proportion of people in the lowest wealth quintiles. (MoHSS and ICF International, 2014).

[4] See Appendix E for a map of Namibia

Only Kwangali, Lozi and Wambo speaking people are included in this sample. The language variable is defined as the language actually spoken by the respondent and not the language which the interview was conducted in. This variable is included to capture any effect "cultural" may have on the probability of smoking, as it may be that one's ethnic group is a determinant in the likelihood of smoking. Just under three quarters of the sample is Wambo, about 16% is Kwangali and just under 10% is Lozi speaking. These figures differ dramatically from those of the Namibian population and the NDHS sample because, of the seven administrative regions which are included in the sample, four are predominantly populated by Wambo speaking people (Ohangwena, Omusati, Oshana and Oshikoto), one is predominantly Lozi speaking (Kavango) and another is mainly Kwangali speaking (Zambezi). The seventh region is the Kunene region. Its population is predominantly Herero speaking, but it is included in the sample because there are Lozi, Kwangali and Wambo speaking respondents interviewed there and it is within 150km of the Angolan and Zambian border.

Education is divided into six categories ranging from respondents who have no education to those who have some higher (post-secondary) education. It appears that a higher proportion of women compared to men have some secondary school education or more, and fewer women have no education in this sample. This is consistent with what is found in the NDHS data as a whole (MoHSS and ICF International, 2014).

Respondents who are married or living with a partner are grouped together. Those who are widowed, divorced, separated/no longer living together or were never in a union are grouped as unmarried. The proportion of married and unmarried respondents in this sample is fairly similar between men and women, however women are less likely to be unmarried than men. This is consistent with what is found in the unrestricted NDHS dataset, however the restricted sample has a higher proportion of unmarried individuals than the NDHS sample as a whole (MoHSS and ICF International, 2014).

The variable for whether one drinks alcohol has been generated from two questions in the data, the first being whether one has ever drunk alcohol before and the second is whether the respondent has drunk alcohol in the last two weeks. Only those who answer yes to the latter question have been classified as drinkers for this analysis, so that only those who can be considered "regular" drinkers are captured as opposed to those who may have only tried alcohol once. According to this definition, it appears that more men are classified as drinkers compared to women. 41.3% of men compared to only 21.7% of women report having had a drink in the two weeks prior to being interviewed.

Religion has been defined into three categories as opposed to the six reported in data. Roman Catholics make up the first category, followed by Protestant/Anglican, Lutheran and Seventh Day Adventists, which have all been grouped together as “Protestants”, in the second category. The third category classifies those who practice “other” religions or those with no religion into one group, to create a new religion variable with only three categories. The largest religious group is Protestant, followed by Roman Catholics, for both men and women.

The variables for the distanced from the borders are divided into three groups depending on how far they live from the border of interest. As previously mentioned only those who live within 150km of either border are included in this sample. As the population of northern Namibia is more concentrated around the Angolan border, most men and women in the sample live within 50km of the Angolan border. This is in stark contrast to the population around the Zambian border which is much smaller. About 90% of both men and women in the sample live more than 50km away from the Zambian border.

**Table 5: Descriptive statistics**[5]

| Variable | | Female (n= 3768 ) | | Male (n= 1676) | |
|---|---|---|---|---|---|
| | | n | Weighted % | n | Weighted % |
| Smokers | | | | | |
| | Non-smokers | 3,727 | 98.95 | 1,471 | 87.82 |
| | Smokers | 40 | 1.05 | 204 | 12.18 |
| Age categories | | | | | |
| | 15-19 | 950 | 25.21 | 478 | 30.31 |
| | 20-24 | 661 | 17.55 | 327 | 20.76 |
| | 25-29 | 581 | 15.43 | 219 | 13.91 |
| | 30-34 | 481 | 12.78 | 158 | 10.04 |
| | 35-39 | 447 | 11.86 | 143 | 9.09 |
| | 40-44 | 358 | 9.49 | 113 | 7.15 |
| | 45-49 | 289 | 7.68 | 78 | 4.97 |
| | 50+ | - | - | 59 | 3.75 |
| Employment status | | | | | |
| | Employed | 1,321 | 35.15 | 773 | 46.22 |
| | Unemployed | 2,437 | 64.85 | 900 | 53.78 |
| Wealth index | | | | | |
| | Poorest | 1,124 | 29.82 | 486 | 28.97 |
| | Poorer | 970 | 25.75 | 434 | 25.87 |
| | Middle | 884 | 23.46 | 416 | 24.79 |
| | Richer | 580 | 15.39 | 251 | 14.99 |
| | Richest | 210 | 5.58 | 90 | 5.38 |
| Rural vs urban | | | | | |
| | Urban | 903 | 23.96 | 1307 | 77.99 |
| | Rural | 2865 | 76.04 | 369 | 22.01 |
| Language of respondent | | | | | |
| | Kwangali | 606 | 16.09 | 259 | 15.48 |
| | Lozi | 362 | 9.61 | 184 | 10.98 |
| | Wambo | 2800 | 74.31 | 1233 | 73.54 |
| Education | | | | | |
| | No education | 150 | 3.99 | 164 | 9.79 |
| | Incomplete primary | 674 | 17.88 | 420 | 25.09 |
| | Complete primary | 288 | 7.64 | 138 | 8.21 |
| | Incomplete secondary | 2010 | 53.34 | 701 | 41.85 |
| | Complete secondary | 466 | 12.37 | 161 | 9.62 |
| | Higher | 180 | 4.78 | 91 | 5.43 |
| Marital status | | | | | |
| | Married/cohabiting | 1062 | 28.18 | 433 | 25.85 |
| | Unmarried | 2706 | 71.82 | 1243 | 74.15 |
| Drinks alcohol | | | | | |

[5] Sample size for each variable may not always add up to that of the sample for men and women, because of missing values for some of the variables.

**Table 5: Descriptive statistics[5]**

| Variable | | Female (n= 3768 ) | | Male (n= 1676) | |
|---|---|---|---|---|---|
| | | n | Weighted % | n | Weighted % |
| | Non-drinker | 2949 | 78.30 | 983 | 58.7 |
| | Drinker | 817 | 21.70 | 692 | 41.3 |
| Religion | | | | | |
| | Roman Catholic | 791 | 21.00 | 456 | 27.19 |
| | Protestant | 2810 | 74.63 | 1105 | 65.96 |
| | other or no religion | 164 | 4.37 | 114 | 6.85 |
| Distance from Angolan border | | | | | |
| | Within 25km | 1600 | 42.45 | 612 | 36.53 |
| | Between 25 and 50km | 1040 | 27.61 | 481 | 28.74 |
| | More than 50km | 1128 | 29.94 | 582 | 34.73 |
| Distance from Zambian border | | | | | |
| | Within 25km | 266 | 7.06 | 135 | 8.04 |
| | Between 25 and 50km | 50 | 1.33 | 37 | 2.20 |
| | More than 50km | 3452 | 91.62 | 1504 | 89.76 |

Source: Namibian Demographic Health Survey, 2013

Table 6 and 7: Smokers by language spoken

Table 6 and 7 show the breakdown of smokers by language group and gender. It is clear from table 6 that a much higher proportion of men smoke than women, as 12.18% of men in the sample smoke compared to only 1.05% of women. This is consistent with findings of other studies done on smoking in Sub-Saharan Africa – smoking prevalence among women is much lower than that of men in general (Townsend, et al., 2006). Lozi men smoke the most, with 32.85% of men in this sample smoking cigarettes, followed by Kwangali men at a rate of 18.73% and Wambo men at 7.71%. Wambo women have the lowest smoking prevalence, with only 0.54% smoking. Kwangali women have the highest prevalence followed by Lozi women at a rate of 3.16% and 1.46%, respectively. These results are in line with those of the NDHS data as a whole, because the prevalence rate of men aged 15 to 49 in the whole sample is just over 18% compared to only 4.2% for women in the same age group.

**Table 6: Male smokers by language spoken**

| | Kwangali | | Lozi | | Wambo | | Total | |
|---|---|---|---|---|---|---|---|---|
| | n=259 | weighted % | n=184 | weighted % | n=1232 | weighted % | n=1675[6] | weighted % |
| Non-smoker | 211 | 81.27 | 124 | 67.15 | 1137 | 92.29 | 1471 | 87.82 |
| Smoker | 49 | 18.73 | 60 | 32.85 | 95 | 7.71 | 204 | 12.18 |

Source: Namibian Demographic Health Survey, 2013

**Table 7: Female smokers by language spoken**

| | Kwangali | | Lozi | | Wambo | | Total | |
|---|---|---|---|---|---|---|---|---|
| | n=606 | weighted % | n=362 | weighted % | n=2799 | weighted % | n=3767 | weighted % |
| Non-smoker | 587 | 96.84 | 357 | 98.54 | 2784 | 99.46 | 3727 | 98.95 |
| Smoker | 19 | 3.16 | 5 | 1.46 | 15 | 0.54 | 40 | 1.05 |

Source: Namibian Demographic Health Survey, 2013

## 3.4 Variable selection

Much has been written about what factors influence the likelihood of smoking or using tobacco products. Since many of the relationships between these factors and tobacco use have been well established in the literature, variables that were found to be statistically significant predictors of tobacco use in other studies have been investigated here to assess whether they should be included as independent variables in this model (Hosmer & Lemenshaw, 2000). The variance inflation factors (VIF) were then calculated to test for multicollinearity. However, multicollinearity was not found to be a problem, as all VIF values were well under 10 (Wooldridge, 2013). These variables were then regressed individually against smoking and a Likelihood Ratio (LR) test was performed each time (Hosmer & Lemenshaw, 2000). The results of these simple logistic regressions were not used for variable selection, but rather to analyse the relationship between the individual independent variables and smoking in this sample, to see if they are in line with what the literature suggests.

### 3.4.1 Literature on variables which influence tobacco use

Studies with a focus on tobacco use in sub-Saharan Africa were used for variables selection. Variables which were found to have a statistically significant effect on the probability of smoking or using tobacco products were included for further analysis. As one will see in the results, these variables are not always significant in the Namibian context.

[6] There is one missing value for male smokers

Some of these studies showed that there is a strong, positive relationship between the probability of smoking, or using other tobacco products, and age (Pampel, 2005; Peer, Bradshaw, Laubscher, & Steyn, 2009). This relationship is often modelled as quadratic by including an age squared term, where probability of smoking is increasing with age until a certain point, after which it begins to decrease (Pampel, 2008; John, Mamudu, & Liber, 2012).

The effect of socioeconomic status (SES) indicators, such as education, employment, and income, on tobacco use in Sub-Sahara Africa is mixed (Townsend, et al., 2006). Results from some studies do seem to indicate a negative relationship between SES indicators and tobacco use. Those who are least educated, have low status employment, and are relatively less wealthy tend to be more likely to smoke tobacco products (Pampel, 2005 & 2008; Peer, Bradshaw, Laubscher, & Steyn, 2009; John, Mamudu, & Liber, 2012).

The effect of rural or urban location on the probability of tobacco use depends on the type of product. Urban dwellers tend to smoke tobacco more than those who live in rural areas, whereas those in rural areas tend to use more smokeless tobacco (Townsend, et al., 2006).

Although the effect of marriage is found to have a statistically significant effect on the probability of using tobacco products by John et al (2012) and Pampel (2005), the direction of the effect differ in these two studies. John et al (2012) find that married Ghanaians tend to be less likely to use tobacco products, whereas Pampel (2005) finds that Malawians and Zambians who are currently married smoke more than those who have never married. John et al (2012) also find that those who report drinking alcohol are much more likely to report using tobacco as well.

In both of Pampel’s studies on tobacco use in Sub-Saharan African countries, he finds that religion has a significant effect on the likelihood smoking. He finds the Protestants are less likely to smoke than Catholics, but those classified as observing “other” religions are most likely to smoke (Pampel, 2005 & 2008).

South African studies on tobacco use tend to control for race (Townsend, Fisher, Gilreath, & King, 2006; Peer, Bradshaw, Laubscher, & Steyn, 2009), but the NDHS does not capture this. Rather, language spoken by the respondent is included to capture the effect of ethnicity.

### 3.4.2 Correlation matrix and contingency tables

Table 3 shows the relationship between the independent variables and dependent variable in a correlation matrix. These correlation coefficients show the strength and direction of the linear relationships between the variables. Contingency tables are also used to assess the relationship between smoking and the categorical independent variables, in Appendix A (Hosmer & Lemenshaw, 2000). The chi-squared hypothesis tests for independence were performed on the contingency tables to see whether the relationships between smoking (the row variable) and the independent variables (the column variables) should be investigated. If the result of the hypothesis test showed an insignificant result, then it could be that the relationship between the row and column variables could be explained by chance (Stockburger, 1996).

The direction of many of the correlation coefficients echo what the literature shows. Age has a positive relationship with smoking. The socioeconomic indicators, show somewhat mixed results, which is also in line with some of the findings in the literature. Being classified as employed has a positive relationship with smoking, meaning that employed men are associated with a higher likelihood of smoking than unemployed men. Wealth index and education have a negative relationship with smoking, which implies that wealthier people and those with more education a lower probability of smoking. Living in a rural area is negatively correlated with smoking, which is expected, as the literature shows this to be the case in other African countries. Being married or cohabiting, and drinking alcohol also have a positive association with smoking.

These results are consistent with the contingency tables. As one can see from appendix A, 19.7% of those who are employed are smokers, as opposed to only 8.36% of those who are classified as unemployed. The chi-squared statistic for this contingency table is significant at the 1% level meaning that there is an association between smoking and employment which is unlikely to be due to chance. The relationship between wealth index and smoking is less clear. Although the poorest group have the highest proportion of smokers, the richest group also has a relatively high proportion. However, the chi-squared test for independence is significant at the 10% level. Marital status, drinking alcohol, language spoken by the respondents and religion all have highly significant tests statistics. 17.76% of married or cohabiting people smoke, compared to only 12.14% of unmarried or non-cohabiting people. Of those who reported drinking in the two weeks prior to the interview, 20.15% are smokers as opposed to only 9.25% of those who did not report drinking. 33.21% of Lozi men and 19.37% of Kwangali men are smokers compared to only 7.69% of Wambo men.

Although it seems that there is a higher proportion of smokers in urban areas, the chi-squared test is not statistically significant with a p-value of 0.115. The contingency table for education and smoking also produces an insignificant result with a p-value of 0.282.

The correlation coefficients and contingency tables of the distance variables seem to tell different stories. While being further away from the Zambian border is negatively associated with smoking, the opposite is true for the Angolan border. The contingency tables for the Angolan border variable shows that those that those who live more than 50km from the Angolan border has the highest proportion of smokers. The positive relationship between being further away from the Angolan border and smoking, is not what the theory would suggest should be the case. Since cigarettes are cheaper in Angola than in Namibia, one would expect that being closer to the border would result in a positive association between smoking and proximity, due to access to cheaper cigarettes. The contingency table for the distance from the Zambian border variable is more in line with what one would expect. 30.39% and 29.79% of those who live within 25km, and 25-50km from the Zambian border, respectively, are smokers, compared to only 10.74% of those who live more than 50km from the border.

It is also important to observe how independent variables are correlated with each other. As one would expect, there is moderate positive relationship between wealth an education, showing that those who are wealthier also tend to have more education. Wealth also has a moderate negative relationship with rural, which implies that those in rural areas are more often classified lower on the wealth index than those in urban areas. There is a weak negative relationship between rural and education, suggesting that men in rural areas tend to be less educated than those in urban areas. Being married or cohabiting shows a strong positive relationship with age, but exhibits a weak positive association with employment. It is expected that those who are older are more likely to be married than younger people, and the same is true for those who are employed albeit to a lesser extent.

## Table 3: Correlation Matrix

| | Smoker | Age | Employed | Wealth index | Rural | Education | married/ cohabiting | Kwangali | Lozi | Wambo | Drinks alcohol | Roman Catholic | Protestant | Other/ no religion | Distance from Angolan border | Distance from Zambian border |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Smoker | 1 | | | | | | | | | | | | | | | |
| Age | 0.0822* | 1 | | | | | | | | | | | | | | |
| Employed | 0.1645* | 0.1976* | 1 | | | | | | | | | | | | | |
| Wealth index | -0.0549 | -0.0278 | 0.0889* | 1 | | | | | | | | | | | | |
| Rural | -0.0385 | -0.0015 | -0.1571* | -0.4627* | 1 | | | | | | | | | | | |
| Education | -0.0191 | -0.1658* | 0.1268* | 0.3609* | -0.2272* | 1 | | | | | | | | | | |
| married/cohabiting | 0.0728* | 0.6293* | 0.2193* | -0.0712* | -0.0412 | -0.0567 | 1 | | | | | | | | | |
| Kwangali | 0.0699* | 0.0309 | 0.0551 | -0.1543* | -0.0451 | -0.0188 | 0.1729 | 1 | | | | | | | | |
| Lozi | 0.2532* | 0.0190 | 0.0877* | -0.0245 | -0.0631* | 0.1274* | 0.1395* | -0.1876* | 1 | | | | | | | |
| Wambo | -0.2561* | -0.0390 | -0.1124* | 0.1383* | 0.0852* | -0.0873* | -0.2445* | -0.6200* | -0.6543* | 1 | | | | | | |
| Drinks alcohol | 0.1558* | 0.1750* | 0.1900* | 0.0671* | -0.0194 | 0.0276 | 0.0215 | -0.1688* | -0.0800* | 0.1939* | 1 | | | | | |
| Roman Catholic | 0.0500 | 0.0043 | 0.0276 | -0.0015 | -0.0386 | -0.0598 | 0.0159 | 0.3020* | -0.1274* | -0.1304* | -0.0225 | 1 | | | | |
| Protestant | -0.1023* | -0.0207 | -0.0456 | 0.0461 | 0.0475 | 0.0704* | -0.0818* | -0.4160* | 0.0015 | 0.3188* | 0.0537 | -0.8312* | 1 | | | |
| Other/no religion | 0.0988* | 0.0295 | 0.0353 | -0.0793* | -0.0206 | -0.0264 | 0.1188* | 0.2399* | 0.2072* | -0.3501* | -0.0582 | -0.1734* | -0.4033* | 1 | | |
| Distance from Angolan border | 0.0856* | 0.0184 | 0.0680* | 0.1085* | -0.0096 | 0.0268 | -0.0042 | -0.3810* | 0.3702* | -0.0024 | 0.0077 | -0.1619* | 0.1759* | -0.0452 | 1 | |
| Distance from Zambian border | -0.1982* | -0.0054 | -0.0630* | -0.0800* | 0.1782* | -0.1492* | -0.1034* | 0.1673* | -0.8508* | 0.5508* | 0.0471 | 0.0853* | 0.0206 | -0.1769* | -0.3231* | 1 |

*p<0.01
Source: Namibia Demographic Health Survey, 2013

### 3.4.3 Simple logistic regressions

Variables which were found to be good predictors of tobacco use and/or smoking were regressed on smoking in separate logistic regressions and LR tests were performed to further analyse the relationship between smoking and the independent variables. These results are presented in table 4 which shows the individual regressions of each independent variable on the binary variable for whether one is a smoker or not. Confidence intervals for each coefficient are displayed in parentheses and the results of each LR test are presented at the bottom of the table. Age and age$^2$ are both highly significant and display a quadratic relationship with age. Employment, marital status, whether one drank alcohol in the two weeks prior to the interview, and language spoken are all highly significant in their respective individual regressions, with p-values of less than 0.01 in the LR tests, meaning that they have significant explanatory power in predicting the likelihood of smoking in this sample.

Only two of the four coefficients on the wealth index are statistically significant at the 10%. According to this simple logistic regression, those who are classified as having a "middle" level of wealth are about 30% (odds ratio 0.69694) less likely to smoke than those classified as being the poorest level of wealth. Those who are classified as being the richest have about 45% (odds ratio 0.54868) lower odds of being smokers than those who are grouped as the poorest.

Although the p-value for the LR test on education is insignificant at 0.302, the odds ratios for having incomplete secondary school or having higher education is significant at the 10% level. Those with incomplete secondary and higher education are associated with having about 33% and 50% lower odds of smoking than those with no education, respectively.

For both the distance variables, only one odds ratio is statistically significant in the individual regressions. Those who live between 25 and 50km from the Angolan border seem to be less likely to smoke than those who live within 25km of it. Although this coefficient is in the expected direction, it is not statistically significant. However, those who live more than 50km from the Angolan border have 1.713 times higher odds of being smokers than those who live within 25km of the Angolan border. This odds ratio to not in the expected direction, but it is significant at the 5% level.

The odds ratios for the regression of distance from the Zambian border on smoking are in the expected direction. Again, those between 25km and 50km of the Zambian border appear to be less likely to smoke than those who live within 25km of the border, but this odds ratio is not significant at all. Men who live more than 50km from the Zambian border, however are 82%

less likely to smoke than those who live within 25km of the Zambian border. This result is in line with what is expected, given what has been said in the literature. Those who live closer to Zambia have access to cheaper cigarettes, therefore the smoking prevalence in these areas may be higher than areas where there is less access to cheaper cigarettes.

The only variables with LR tests which are not statistically significant at even the 10% level, are education and location (rural vs urban). This is an unexpected result for education, considering that some of its odd ratios are significant and there is substantial evidence in the literature showing that there is a relationship between education and the likelihood of smoking.

As one can see in table 3, there is a fairly strong correlation between education and the wealth index. It may be that education does not have a significant impact on the probability of being a smoker on its own, but when included in a regression with other socioeconomic indicators, may be jointly significant. The full model was run with and without education, and another LR test was performed to see if the explanatory power of the model is improved by the inclusion of the education variable. Again, the p-value of 0.1426 from the LR test is insignificant. However, when both education and wealth are removed from the model and the LR test is performed it results in a p-value of 0.0947 ($\chi^2_{10} = 16.18$), meaning that at the 10% level we reject the hypothesis that both education and wealth have no explanatory power in this model. This shows that although the simple logistic regression shows that education does not have a significant effect on the odds of being a smoker, its correlation with wealth means that its exclusion may introduce endogeneity into the model, resulting in biased estimates.

**Table 4: Preliminary regressions**

| Independent variables | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Age | 1.36430*** (1.26258 - 1.47422) | | | | | | | | | | |
| Age2 | 0.99585*** (0.99473 - 0.99697) | | | | | | | | | | |
| Employment (employed=1; unemployed=0) | | 2.68784*** (1.99954 - 3.61308) | | | | | | | | | |
| Poorer (poorest is the reference group) | | | 0.91148 (0.63756 - 1.30308) | | | | | | | | |
| Middle | | | 0.69694* (0.47439 - 1.02388) | | | | | | | | |
| Richer | | | 0.54868** (0.33820 - 0.89016) | | | | | | | | |
| Richest | | | 0.89161 (0.47248 - 1.68255) | | | | | | | | |
| Rural (rural=1, urban=0) | | | | 0.77010 (0.55599 - 1.06667) | | | | | | | |
| Incomplete primary (no education is the reference group) | | | | | 0.67341 (0.41124 - 1.10271) | | | | | | |
| Complete primary | | | | | 0.71244 (0.37602 - 1.34985) | | | | | | |
| Incomplete secondary | | | | | 0.66629* | | | | | | |

**Table 4: Preliminary regressions**

| Independent variables | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | (0.42087 - 1.05482) | | | | | | |
| Complete secondary | | | | | 0.94828 (0.54037 - 1.66411) | | | | | | |
| Higher | | | | | 0.49958* (0.22515 - 1.10853) | | | | | | |
| Married/cohabiting (married/cohabiting =1; single=0) | | | | | | 1.56308*** (1.16339 - 2.10008) | | | | | |
| Drank alcohol in the last two weeks (drinker =1; non-drinker = 0) | | | | | | | 2.47640*** (1.86262 - 3.29243) | | | | |
| Protestant (Catholics are the reference group) | | | | | | | | 0.63271*** (0.46249 - 0.86558) | | | |
| Other or no religion | | | | | | | | 1.71501** (1.07398 - 2.73867) | | | |
| Distance from Angolan border: between 25 and 50km (0-25km is the reference group) | | | | | | | | | 0.87133 (0.58063 - 1.30756) | | |
| Distance from Angolan border: More than 50km | | | | | | | | | 1.71338*** (1.23298 - 2.38096) | | |
| Distance from Zambian border: between 25 and 50km (0-25km is the reference group) | | | | | | | | | | 0.97165 | |

**Table 4: Preliminary regressions**

| Independent variables | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | (0.48605 - 1.94240) | |
| Distance from Zambian border: More than 50km | | | | | | | | | | 0.27570*** (0.19579 - 0.38823) | |
| Lozi (Kwangali is the reference group) | | | | | | | | | | | 2.07038*** (1.38826 - 3.08767) |
| Wambo | | | | | | | | | | | 0.34661*** (0.23700 - 0.50692) |
| Constant | 0.00102*** (0.00029 - 0.00358) | 0.09125*** (0.07192 - 0.11576) | 0.19185*** (0.15103 - 0.24370) | 0.19388*** (0.14599 - 0.25746) | 0.21970*** (0.14698 - 0.32839) | 0.13819*** (0.11637 - 0.16410) | 0.10188*** (0.08221 - 0.12626) | 0.19837*** (0.15432 - 0.25500) | 0.12784*** (0.09814 - 0.16652) | 0.43662*** (0.32399 - 0.58840) | 0.24020*** (0.17585 - 0.32808) |
| Observations | 1,675 | 1,672 | 1,675 | 1,675 | 1,675 | 1,675 | 1,675 | 1,674 | 1,675 | 1,675 | 1,675 |
| Pseudo R-squared | 0.0617 | 0.0343 | 0.00618 | 0.00179 | 0.00452 | 0.00636 | 0.0298 | 0.0163 | 0.0135 | 0.0427 | 0.0860 |
| LR test Chi squared statistic | 82.42 | 45.79 | 8.263 | 2.394 | 6.047 | 8.496 | 39.88 | 21.77 | 17.98 | 57.12 | 115 |
| Prob > Chi2 (p-value from LR test) | 0.000 | 0.000 | 0.082 | 0.122 | 0.302 | 0.004 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |

Confidence intervals in parentheses

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

### 3.5 The model

A multivariable logistic regression is used to determine the probability of the smoking, given the distance one lives from the northern Namibian borders and other characteristics.

$$P(Smoker=1|\boldsymbol{X}) = \Lambda(\beta_0 + \boldsymbol{\beta X})$$

Where the dependent variable is a binary, with an outcome equal to one if a respondent is a smoker and zero otherwise. ***X*** represents the vector of independent variables which include age, $age^2$, employment status, wealth index, rural/urban location, language spoken by respondent, education, marital status, drinks alcohol, religion, distance from the Angolan border, and distance from the Zambian border. These variables, besides those for distances from the borders, were chosen based on their ability to predict the likelihood of smoking. In theory, once these variables have been controlled for, it would be clear whether proximity to the border has an effect on smoking prevalence.

# 4. Results

The following tables present the results of the logistic regressions. For tables 8 to 11, column 1 of each table contains the restricted model, which excludes the distance variables and column 2 contains the unrestricted model, which includes the distance variables. As most of the variables in these regressions are categorical, results are reported as odds ratios, since partial effect at the average and average partial effects are not straight forward to interpret for categorical variables. The odds ratio is defined as

$$Odds\ ratio = e^{\beta i}$$

Where $\beta_i$ is coefficient on $x_i$ in the logit regression. The odds ratio is interpreted as how much more likely it is that the outcome of interest would be present when $x$=1, as opposed to when $x$=0 (Hosmer & Lemenshaw, 2000). In other words odds ratios tell us how much more likely one is to be a smoker when, for instance, one drinks alcohol (*alcohol*=1), as opposed to when one does not drink (*alcohol*=0), when all other factors are held constant.

The 95% confidence intervals are reported in parentheses below the odds ratios. The interpretation of the confidence interval is that, if one were to continuously resample from the same population and run this same regression every time on each of these samples, the true population odds ratio would lie within this confidence interval 95% of the time (Wooldridge, 2013). If the confidence interval contains one, then we fail to reject the null hypothesis that the true population odds ratio is equal to one (UCLA: Statistical Consulting Group, n.d. a), meaning that the variable of interest has no effect on probability of smoking, all other factors held constant.

In all the regressions the base categories for the independent categorical variables are, "poorest" for the wealth index, "no education" for the education quintiles, Kwangali speakers for the languages, Roman Catholic for the religion categories, and residing between 0 and 25km from the border for both distance variables.

## 4.1 Final model

Table 8 shows the results for logistic regressions for Wambo, Lozi and Kwangali men. The coefficients on age and $age^2$ are significant at the 1% level, which is in line with the findings of the literature, that smoking prevalence increases at a decreasing rate with age (Pampel, 2008). In the restricted model, the likelihood of being a smoker is increasing in age until 38.87 years after which it is decreasing, when all other variables are held constant. The addition of the distance variables only increases the turning point of the model slightly to 38.94 years.

The binary variable for employment is also highly significant. In the logistic regression without the distance variables, the odds of smoking are higher if one is employed, as opposed to unemployed (odds ratio 1.73787). However, these odds decrease when the distance variables are added (odds ratio 1.67521), but remain highly significant at the 1% level. This means that employed men are 1.68 times more likely to be smokers than unemployed men, all other factors held constant. The other socioeconomic variables, such as wealth index and education, show almost no significance at all.

The wealth index shows that those in the "poorer" category (the second lowest quintile) have slightly higher odds of smoking than the "poorest" (the lowest quintile), whereas the "middle", "richer" and "richest" groups are less likely to smoke than the poorest. Since these coefficients are insignificant, not much can be extrapolated from this result, except that living standards do not appear to have any effect on the likelihood of smoking.

The education variable shows that the probability of being a smoker decreases, as the level of educations attained increases. Only the odds ratios on "completed secondary" and "higher" education are significant. Those with completed secondary education are about half as likely to smoke as those with no education at all, with odds ratios of 0.52109 and 0.50688 in the restricted and unrestricted models respectively. The addition of the distance variables reduces the magnitude of the odds ratio slightly, but increases the significance of the coefficients from the 10% to the 5% level of significance. The coefficients on the higher education are also significant at the 5%, but do not change much with the addition of distance variables. Those with higher education have about 68% lower odds of smoking than those with no education.

Although it appears that those in rural areas are less likely to be smokers than those in urban areas, location does not seem to have a statistically significant effect on the likelihood of the men in this sample being cigarette smokers.

The logistic regressions show that marriage and cohabiting have a statistically significant effect on the probability of being a smoker. Adding distance variables reduces the size of the coefficients slightly to make the odds of smoking for married or cohabiting men 38.2% (odds ratio 0.61832) lower than their single counterparts.

The odds ratios for language of the respondent are highly significant in all the regressions. Of the three ethnic groups represented in these results, Lozi men are the most likely to smoke. Adding distance variables to the logistic regression increases the size of the coefficient significantly from an odds ratio of 2.23394 to 3.06910, meaning that in the unrestricted model, Lozi men are 3 times as likely to smoke as Kwangali men. Wambo men are the least likely to smoke of the three groups, as they have about 78% (odds ratio 0.22096) lower odds of smoking than Kwangali men according to the unrestricted regression.

There is a strong positive relationship between being a drinker and smoking among men in this sample. Those who report drinking alcohol in the two weeks preceding the survey are more than three times more likely to be smokers than those do not report drinking in the last two weeks, with odds ratios of 3.35725 and 3.47977 in the restricted and unrestricted models respectively. The coefficients are significant at the 1% level in all regressions.

When distance variables are added to the logistic regression, the odds ratio for "Protestant" for the religion variable is significant at the 10%. This suggests that Protestants are about 28% (odds ratio 0.72107) less likely to be smokers than Roman Catholics. There does not appear to be a statistically significant difference in smoking rates between Catholics and those who are classified as "other or no religion".

It appears that the distance from the Angolan and Zambian border has no effect on the likelihood of smoking cigarettes. All distance categories, except for one, are not statistically significant at all. Those who live between 25 and 50km away from the Angolan border are about 49% (odds ratio 1.49021) more likely to smoke than those who live with 25km of the border. This effect appears to be driven by Kwangali men, as this result only appears again in the regression for Kwangali men (see table 10).

**Table 8: Regression results of all men**

| Independent variables | (1) | (2) |
|---|---|---|
| Age | 1.26318*** | 1.26037*** |
| | (1.15680 - 1.37934) | (1.15398 - 1.37657) |
| Age2 | 0.99700*** | 0.99703*** |
| | (0.99578 - 0.99822) | (0.99581 - 0.99826) |
| Employed (employed=1; unemployed=0) | 1.73787*** | 1.67521*** |
| | (1.23761 - 2.44035) | (1.18867 - 2.36091) |
| Poorer (reference group is poorest) | 1.17952 | 1.17197 |
| | (0.78148 - 1.78030) | (0.77285 - 1.77722) |
| Middle | 0.81747 | 0.81178 |
| | (0.50818 - 1.31500) | (0.50197 - 1.31279) |
| Richer | 0.71017 | 0.68496 |
| | (0.38192 - 1.32054) | (0.36474 - 1.28628) |
| Richest | 0.85818 | 0.84200 |
| | (0.36222 - 2.03321) | (0.35228 - 2.01250) |
| Rural (rural=1, urban=0) | 0.74075 | 0.69985 |
| | (0.47923 - 1.14498) | (0.44948 - 1.08967) |
| Incomplete primary (reference group is no education) | 0.83605 | 0.81231 |
| | (0.47881 - 1.45984) | (0.46321 - 1.42452) |
| Complete primary | 0.89671 | 0.87927 |
| | (0.43788 - 1.83633) | (0.42791 - 1.80670) |
| Incomplete secondary | 0.69864 | 0.69105 |
| | (0.40913 - 1.19302) | (0.40208 - 1.18770) |

**Table 8: Regression results of all men**

| Independent variables | (1) | (2) |
|---|---|---|
| Complete secondary | 0.52109* | 0.50688** |
| | (0.26971 - 1.00679) | (0.26020 - 0.98743) |
| Higher | 0.31292** | 0.31922** |
| | (0.11948 - 0.81953) | (0.12129 - 0.84015) |
| Married/cohabiting (married = 1; unmarried=0) | 0.62009** | 0.61832** |
| | (0.41075 - 0.93612) | (0.40859 - 0.93570) |
| Lozi (reference group is Kwangali) | 2.23394*** | 3.06910*** |
| | (1.39096 - 3.58781) | (1.32876 - 7.08885) |
| Wambo | 0.24978*** | 0.22096*** |
| | (0.15364 - 0.40607) | (0.13133 - 0.37176) |
| Drank alcohol in the last two weeks (drank=1; did not drink=0) | 3.35725*** | 3.47977*** |
| | (2.38118 - 4.73342) | (2.45696 - 4.92837) |
| Protestant (reference group is Catholic) | 0.73664 | 0.72107* |
| | (0.50596 - 1.07251) | (0.49376 - 1.05304) |
| Other or no religion | 0.82178 | 0.87409 |
| | (0.47537 - 1.42062) | (0.50204 - 1.52187) |
| Distance from the Angolan border: between 25 and 50km (reference group is within 25km) | | 1.49021* |
| | | (0.93431 - 2.37685) |
| Distance from the Angolan border: more than 50km | | 1.29418 |
| | | (0.81596 - 2.05267) |
| Distance from the Zambian border: between 25 and 50km (reference group is within 25km) | | 0.77844 |
| | | (0.34928 - 1.73493) |
| Distance from the Zambian border: more than 50km | | 1.64080 |
| | | (0.78725 - 3.41980) |
| Constant | 0.00584*** | 0.00359*** |
| | (0.00118 - 0.02886) | (0.00063 - 0.02046) |
| Observations | 1,671 | 1,671 |

**Table 8: Regression results of all men**

| Independent variables | (1) | (2) |
|---|---|---|
| Pseudo R-squared | 0.198 | 0.202 |

Confidence intervals in parentheses

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

Source: Namibian Demographic Health Survey, 2013

The following three tables 9, 10 and 11 show the regression results for Wambo, Lozi and Kwangali men, respectively. Separate regressions were run for men of different ethnic groups to see whether the "border effect" would be different for each group. Table 6 shows that there is a significant difference in smoking rates between the three ethnic groups and the regression results presented in table 8 also show that there is a statistically significant difference in the probability of smoking between them. It may be that these differences are driven by a "cultural effect", as the social norms around smoking for different ethnicities may play a role. Given these differences, separate logistic regressions are run for Wambo, Kwangali and Lozi men. The same explanatory variables included in the regressions for all men have been included in the regressions for the separate ethnic groups.

Table 9 shows the logistic regression results for Wambo men. The results are very similar to those in table 8. Many of the variables which are significant in table 8 are significant in table 9 and the direction of their effects are similar. This is most likely because, of the sample of 1671 men in the regression with all men, 1144 are Wambo, therefore much of what we observe in table 8 is driven by the behaviour of Wambo men. Again, the inclusion of the distance variable changed the magnitude of the odds ratios slightly, but had little effect of their significance or the direction of their effects.

As with the regression in table 8, age has a quadratic, and employment has a positive relationship with the likelihood of being a smoker in both the regressions. Both these variables are highly significant. The probability of being a smoker increases with age until about 45 years in both regressions, after which it declines. Those who are employed are about 1.8 times more likely to be smokers than men classified as unemployed, in both models. As with table 8, the wealth indices and location are still statistically insignificant.

Unlike table 8, only those who have incomplete and complete secondary are found to be less likely to smoke than those with no education, at a statistically significant level in the logistic regressions. In both the restricted (odds ratio 0.47790) and unrestricted (odds ratio 0.47544) logistic regressions, those with incomplete secondary education are about half as likely to be smokers compared to those with no education. Those with complete secondary education have 65.6% (odds ratio 0.34393) lower odds of smoking than those with no education, when the distance variable is added.

Wambo men who reported drinking alcohol in the two weeks before the survey are 1.62624 times more likely to be smokers, than those who did not report drinking in over this period. The odds ratio is slightly higher (odds ratio 1.63108) when the distance categories are added.

As with the previous set of results, those who are married or cohabiting are still found to be far less likely to smoke than those who are not, however none of the coefficients on the religion variable are statistically significant for Wambo men.

The real variable of interest here is the distance from the Angolan border. Since Wambo people in northern Namibia mainly live in the regions neighbouring the Angolan border, only the distance from the Angolan border has been considered in this regression. The odds ratios in the logistic regression are slightly greater than one, which suggests that those further from the border are more likely to smoke than those within 25km of the Angolan border. However, this variable is completely statistically insignificant in all regressions for Wambo men, meaning that the distance from the Angolan border actually has no effect on the likelihood of Wambo men being smokers.

**Table 9: Regression results of Wambo men**

| Independent variables | (1) | (2) |
|---|---|---|
| Age | 1.24896*** | 1.24857*** |
| | (1.11119 - 1.40382) | (1.11061 - 1.40367) |
| $Age^2$ | 0.99753*** | 0.99753*** |
| | (0.99596 - 0.99911) | (0.99596 - 0.99911) |
| Employed (employed=1; unemployed=0) | 1.80541** | 1.80316** |
| | (1.07354 - 3.03622) | (1.06909 - 3.04128) |
| Poorer (reference group is poorest) | 1.22624 | 1.21986 |
| | (0.65536 - 2.29441) | (0.65084 - 2.28638) |
| Middle | 0.82495 | 0.81670 |
| | (0.41228 - 1.65068) | (0.40608 - 1.64257) |
| Richer | 0.73058 | 0.72154 |
| | (0.32046 - 1.66557) | (0.31405 - 1.65774) |
| Richest | 2.03222 | 2.01496 |
| | (0.71450 - 5.78018) | (0.70613 - 5.74975) |
| Rural (rural=1, urban=0) | 0.80083 | 0.80094 |
| | (0.43729 - 1.46659) | (0.43499 - 1.47475) |
| Incomplete primary (reference group is no education) | 0.63847 | 0.63570 |
| | (0.31614 - 1.28942) | (0.31437 - 1.28549) |
| Complete primary | 0.83444 | 0.83105 |
| | (0.33373 - 2.08638) | (0.33193 - 2.08072) |
| Incomplete secondary | 0.47790** | 0.47544** |
| | (0.23417 - 0.97533) | (0.23170 - 0.97558) |
| Complete secondary | 0.34822* | 0.34393* |
| | (0.11894 - 1.01954) | (0.11634 - 1.01675) |
| Higher | 0.40572 | 0.40185 |
| | (0.12461 - 1.32101) | (0.12250 - 1.31822) |
| Married (married=1; unmarried=0) | 0.37439*** | 0.37420*** |
| | (0.19096 - 0.73400) | (0.19072 - 0.73419) |
| Drank alcohol in the last two weeks (drank=1; did not drink=0) | 1.62624** | 1.63108** |

**Table 9: Regression results of Wambo men**

| Independent variables | (1) | (2) |
|---|---|---|
| | (1.00182 - 2.63985) | (1.00303 - 2.65236) |
| Protestant (reference group is Catholic) | 0.80530 | 0.80376 |
| | (0.47744 - 1.35832) | (0.47625 - 1.35649) |
| Other or no religion | 2.60386 | 2.62210 |
| | (0.64055 - 10.58469) | (0.63705 - 10.79257) |
| Distance from the Angolan border: between 25 and 50km (reference group is within 25km) | | 1.07316 |
| | | (0.59049 - 1.95037) |
| Distance from the Angolan border: more than 50km | | 1.03817 |
| | | (0.58831 - 1.83202) |
| Constant | 0.00204*** | 0.00200*** |
| | (0.00022 - 0.01853) | (0.00022 - 0.01828) |
| | | |
| Observations | 1,144 | 1,144 |
| Pseudo R-squared | 0.113 | 0.113 |

Confidence intervals in parentheses

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

Source: Namibian Demographic Health Survey, 2013

Table 10 and 11 show the results of the regressions for Lozi and Kwangali men, respectively. Of the three ethnic groups, Lozi and Kwangali men have the highest proportion of smokers, as shown in table 6. As with the previous set of results, the inclusion of the distance variables causes only a small change in the magnitude of odds ratios and no change in their signs, in either of the regressions. In both regressions age is statistically significant and displays a quadratic relationship with smoking. However age and $age^2$ lose their significance slightly, going from being significant at the 1% level to only the 5% level, when the distance variables are added to the regression for Kwangali men. Among Lozi men, the likelihood of smoking increases with age until about 34.6 years, after which it starts to decline. The same is true from Kwangali men, but the turning point is slightly higher, at 35 years of age.

Lozi men who are employed are more than twice as likely to be smokers as those who are not. When the distance variable is added to the regression, the odds ratio falls from 2.30864 to

2.19018, but it remains significant at the 5% level. However, among Kwangali men, employment is found to have no statistically significant relationship with the likelihood of smoking.

Lozi men who are higher on the wealth index appear to be less likely to smoke cigarettes than those who are ranked poorest. Unlike with Wambo men, some of the odds ratios for the wealth index variable show a statistically significant effect on smoking, even if only at the 10% level. Those who fall into the richer category are about 79% less likely to be smokers than those ranked poorest, with odds ratios of 0.21065 and 0.20716 in the restricted and unrestricted models, respectively. Lozi men who are in the richest category have 82.1% lower odds of smoking cigarettes than those in the poorest category, with odds ratio of 0.17934 in the unrestricted model. Unlike the regressions for Lozi and Wambo men, the odds ratios on the wealth index for Kwangali men are not consistently in one direction. However these coefficients are mostly statistically insignificant. Only the coefficient on the poorer wealth group is significant (at the 10% level) for Kwangali men, showing that those who are "poorer" are more than twice as likely to smoke as those classified as "poorest", with odds ratios of 2.29912 and 2.42475 for the restricted and unrestricted models, respectively.

As with Wambo men, indicating that one had had alcohol in the past two weeks seems to be a major predictor for whether one is a cigarette smoker for both Lozi and Kwangali men. Among Lozi men, the addition of the distance variables amplifies this effect by increasing the odds ratio from 5.63125 to 5.77643. According to the unrestricted regression, those who reported drinking two weeks before being interviewed are almost 6 times more likely to be smokers than those who did not. This variable is significant at the 1% level. For Kwangali men the effect is even stronger. Kwangali men who report having had alcohol within two weeks of being interviewed are about ten times as likely to report smoking cigarettes as those who have not reported drinking in within two weeks of the interview. The size of the odds ratio decreases when distance variables are added from 10.08332 to 9.79611.

Religion and education appear to have no statistically significant association with the odds of smoking for Lozi men. However for Kwangali men, those who have complete primary education are associated with having almost four times higher odds of smoking than those with no education at all, with an odds ratio of 3.83719 in the unrestricted model. This odds ratio is significant at the 10% level, but none of the other odds ratios for this variable are statistically significant. The effect of religion on the odds of smoking is significant at the 10% level in the restricted regression for Kwangali men, as Protestants are found to have about 61% lower odds of smoking than Catholics (odds ratio 0.39363). The addition of the distance variable, however, make the effect of religion statistically insignificant although the direction of the effect remains unchanged.

Location and marriage/cohabiting however, appear to have no statistically significant effect on the probability of smoking among Lozi or Kwangali men.

Again, the distance variable is of primary interest. Lozi people in north-eastern Namibia primarily live in the Zambezi region which only borders Zambia, as a result only the distance from the Zambian border is considered in this set of regressions. The coefficient for those who live 25km to 50km from the Zambian border indicated that they are 31.3% (odds ratio 0.68657) less likely to smoke than those who live within 25km of the border, whereas those who live more than 50km away are almost 18% more likely to smoke (odds ratio 1.17941). However, these coefficients are not statistically significant at all, meaning that the distance from the Zambian border has no effect on the odds of smoking for Lozi men.

Distance from the border seems to only have an effect on Kwangali men, however the direction of the effect is not what one would expect. Living between 25km and 50km of the Angolan border has a strong positive effect on the odds of smoking. Kwangali men who live within this distance of the border are more than three time as likely to smoke as those live within 25km of the Angolan border (odds ratio 3.17259). Living more than 50km from the border has no statistically significant effect on the odds of smoking.

**Table 10: Regression results for Lozi men**

| Independent variables | (1) | (2) |
|---|---|---|
| Age | 1.41011*** | 1.40673*** |
| | (1.12890 - 1.76137) | (1.12605 - 1.75737) |
| Age2 | 0.99506*** | 0.99508*** |
| | (0.99188 - 0.99825) | (0.99191 - 0.99827) |
| Employed (employed=1; unemployed=0) | 2.30864** | 2.19018** |
| | (1.17446 - 4.53811) | (1.09734 - 4.37138) |
| Poorer (reference group is poorest) | 0.60328 | 0.60418 |
| | (0.27079 - 1.34400) | (0.26765 - 1.36387) |
| Middle | 0.69087 | 0.67929 |
| | (0.28214 - 1.69170) | (0.27192 - 1.69693) |
| Richer | 0.21065** | 0.20716** |
| | (0.04605 - 0.96363) | (0.04468 - 0.96046) |
| Richest | 0.18225* | 0.17934* |
| | (0.02767 - 1.20039) | (0.02717 - 1.18387) |
| Rural (rural=1, urban=0) | 0.50549 | 0.52229 |
| | (0.18005 - 1.41920) | (0.18067 - 1.50980) |
| Incomplete primary (reference group is no education) | 1.11989 | 1.03732 |
| | (0.27469 - 4.56566) | (0.24774 - 4.34341) |
| Complete primary | 0.37571 | 0.37407 |
| | (0.05386 - 2.62076) | (0.05387 - 2.59772) |
| Incomplete secondary | 1.16860 | 1.11458 |
| | (0.30423 - 4.48882) | (0.28555 - 4.35057) |
| Complete secondary | 0.91800 | 0.86921 |
| | (0.22554 - 3.73649) | (0.20889 - 3.61680) |
| Higher | 0.27828 | 0.27516 |
| | (0.02662 - 2.90880) | (0.02593 - 2.91974) |
| Married/cohabiting (married=1; unmarried=0) | 0.57249 | 0.58712 |
| | (0.26541 - 1.23488) | (0.27069 - 1.27344) |
| Drank alcohol in the last two weeks (drank=1; did not drink=0) | 5.63125*** | 5.77643*** |

**Table 10: Regression results for Lozi men**

| Independent variables | (1) | (2) |
|---|---|---|
| | (2.96674 - 10.68883) | (3.01850 - 11.05422) |
| Protestant (reference group is Catholic) | 0.98705 | 0.96833 |
| | (0.41211 - 2.36413) | (0.39784 - 2.35686) |
| Other or no religion | 0.85326 | 0.90217 |
| | (0.30957 - 2.35179) | (0.32328 - 2.51764) |
| Distance from the Zambian border: between 25 and 50km (reference group is with 25km) | | 0.68657 |
| | | (0.27942 - 1.68697) |
| Distance from the Zambian border: more than 50km | | 1.17941 |
| | | (0.46684 - 2.97961) |
| Constant | 0.00269*** | 0.00304*** |
| | (0.00005 - 0.13683) | (0.00006 - 0.15749) |
| Observations | 276 | 276 |
| Pseudo R-squared | 0.223 | 0.226 |

Confidence intervals in parentheses

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

Source: Namibian Demographic Health Survey, 2013

**Table 11: Regression results for Kwangali men**

| Independent variables | (1) | (2) |
|---|---|---|
| Age | 1.37032*** | 1.35213** |
| | (1.08240 - 1.73483) | (1.06375 - 1.71868) |
| Age2 | 0.99550*** | 0.99566** |
| | (0.99216 - 0.99885) | (0.99226 - 0.99907) |
| Employment (employed=1; unemployed=0) | 1.42493 | 1.60628 |
| | (0.60676 - 3.34634) | (0.66722 - 3.86701) |
| Poorer (reference group is poorest) | 2.29912* | 2.42475* |
| | (0.92785 - 5.69703) | (0.91583 - 6.41975) |
| Middle | 0.59637 | 0.56159 |
| | (0.13450 - 2.64423) | (0.12468 - 2.52968) |
| Richer | 1.15652 | 1.08272 |
| | (0.21561 - 6.20345) | (0.18255 - 6.42174) |
| Richest | 0.32929 | 0.33802 |
| | (0.01729 - 6.27062) | (0.01566 - 7.29488) |
| Rural (rural=1, urban=0) | 0.61428 | 0.41741 |
| | (0.20190 - 1.86891) | (0.12630 - 1.37954) |
| Incomplete primary (reference group is no education) | 1.60372 | 1.38737 |
| | (0.40167 - 6.40295) | (0.33951 - 5.66926) |
| Complete primary | 3.90383* | 3.83719* |
| | (0.80688 - 18.88747) | (0.77676 - 18.95571) |
| Incomplete secondary | 1.11294 | 0.98409 |
| | (0.30916 - 4.00651) | (0.26300 - 3.68218) |
| Complete secondary | 0.55372 | 0.56709 |
| | (0.09312 - 3.29256) | (0.09238 - 3.48098) |
| Higher | 0.54192 | 0.50309 |
| | (0.05368 - 5.47060) | (0.04417 - 5.73070) |
| Married/cohabiting (married = 1; unmarried=0) | 0.90738 | 0.90246 |
| | (0.35823 - 2.29834) | (0.33901 - 2.40239) |
| Drank alcohol in the last two weeks (drank=1; did not drink=0) | 10.08332*** | 9.79611*** |

**Table 11: Regression results for Kwangali men**

| Independent variables | (1) | (2) |
|---|---|---|
| | (4.20461 - 24.18143) | (3.98231 - 24.09750) |
| Protestant (reference group is Catholic) | 0.39363* | 0.48653 |
| | (0.13670 - 1.13348) | (0.15703 - 1.50741) |
| Other or no religion | 0.57981 | 0.61187 |
| | (0.21259 - 1.58132) | (0.22180 - 1.68797) |
| Distance from the Angolan border: between 25 and 50km (reference group is within 25km) | | 3.17259** |
| | | (1.18533 - 8.49161) |
| Distance from the Angolan border: more than 50km | | 0.68703 |
| | | (0.13035 - 3.62119) |
| Constant | 0.00102*** | 0.00137*** |
| | (0.00002 - 0.06583) | (0.00002 - 0.09375) |
| | | |
| Observations | 251 | 251 |
| Pseudo R-squared | 0.233 | 0.257 |

Confidence intervals in parentheses

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

Source: Namibian Demographic Health Survey, 2013

## 4.2 Likelihood ratio tests

The likelihood ratio (LR) test was used on each logistic regression to assess whether adding the distance variable increased the explanatory power of the model. Each LR test, except for that on Kwangali men, produced small chi-squared ($\chi^2$) values and therefore very large p-values. In each of the LR tests the unrestricted model was identical to the restricted one, except for the addition of the distance variables. For the regression which included all men chi-squared statistic for the LR test was 5.47 with a p-value of 0.2422. The LR tests on the logistic regressions for Wambo, Lozi and Kwangali men separately produced p-values of 0.9735 ($\chi^2$ = 0.05), 0.5922 ($\chi^2$ = 1.05) and 0.0494 ($\chi^2$ = 6.02) respectively. Only the LR test for Kwangali men is statistically significant at the 5% level, meaning that the addition of the distance variable to the regression for Kwangali men increased the explanatory power of the model at a statistically significant level. This, however is not the case for any of the other regressions. The inclusion of the distance variables in the regressions for all men, Wambo and Lozi men did not help to explain likelihood of being a smoker.

### 4.3 Smoking intensity

Additional logistic regressions were run to see whether the inclusion of the distance variables had an impact on the intensity of smoking, even if they have no statistically significant impact on the probability of being a smoker. The dependent variable, smoking intensity, was derived from the question, “How many cigarettes have you smoked in the last 24 hours?” which was only asked to those who were already smokers. There is no standard way of defining what is considered heavy smoking as opposed to light or moderate smoking (Husten, 2009). Therefore, for the purposes of this study, men who reported smoking between 0 and 5 cigarettes in the 24 hours before the interview are classified as light smokers, those who smoked between 6 or more cigarettes are heavy smokers. The area under consideration, those with 150km of the northern borders, is rural and relatively poor compared to other parts of the country, therefore six or more cigarettes is classified as heavy smoking, as smoking more than this would be fairly expensive. Table 12 shows that of the 224 smokers in our sample, that vast majority are light smokers, as 80.87% are classified as light smokers as opposed to 19.13 who are heavy smokers.

**Table 12: Proportion of heavy vs light smokers**

| | n=224 | weighted % |
|---|---|---|
| Light smokers | 181 | 80.87 |
| Heavy Smokers | 43 | 19.13 |

Source: Namibian Demographic Health Survey, 2013

Table 13 presents the results for this regressions. Column 1 contains the results for the restricted regression with no distance variables, whereas column two shows the results for the unrestricted regression which includes the distance variables. As there are only 224 observations in these regressions, the contingency tables for the distance variables and smoking contained some cells with no observations. To correct for this, the distance variables were converted into binary variables, which are coded as one for those who live within 25km of the border of interest and zero otherwise. Odds ratios are reported for each variable, along with a 95% confidence interval.

The inclusion of the distance variables does not change the results of the regression by much. Variables which were statistically insignificant in the restricted regression are still insignificant, and those which were statistically significant retain their significance. The results of the unrestricted regression show that those in the richest wealth quintile are associated with having 15.641 times higher odds of being a high intensity smokers as opposed to low intensity smokers, compared to those who are classified as the poorest, when all other factors are held constant. This odds ratio increased from 12.941 to 15.641, when distance the distance variables were added to the model. Those with incomplete primary education are 7.217 times more likely to be high intensity smokers, as opposed to the low intensity smokers, compared to those with no education. Oshiwambo speaking men are about 85% less likely to be high intensity smokers, as opposed to low intensity smokers than Kwangali men. Men who reported drinking alcohol in the two weeks prior to the interview are just over twice as likely to be high intensity smokers, than low intensity smokers, compared to those who did not report drinking. Age, age$^2$, employment, location, marriage/cohabiting and religion are all have a statistically insignificant effect on smoking intensity in this regression.

The results of this regression show that one's distance from the northern borders does not have a statistically significant effect on smoking intensity, either. Given this finding, it is no surprise that the LR test, to assess whether the addition of the distance variables adds to the predictive power of the model, also produced a highly insignificant p-value of 0.2969 ($\chi^2_2 = 2.43$).

**Table 13: Smoking intensity**

| Independent variables | (1) | (2) |
|---|---|---|
| Age | 1.148 | 1.179 |
| | (0.901 - 1.461) | (0.920 - 1.511) |
| $Age^2$ | 0.999 | 0.998 |
| | (0.995 - 1.002) | (0.995 - 1.002) |
| Employment (employed=1; unemployed=0) | 1.356 | 1.221 |
| | (0.570 - 3.226) | (0.507 - 2.942) |
| Poorer (reference group is poorest) | 0.898 | 1.032 |
| | (0.296 - 2.720) | (0.328 - 3.252) |
| Middle | 2.165 | 2.559 |
| | (0.679 - 6.899) | (0.782 - 8.371) |
| Richer | 3.060 | 3.530 |
| | (0.696 - 13.452) | (0.766 - 16.271) |
| Richest | 12.941** | 15.641*** |
| | (1.653 - 101.341) | (1.951 - 125.391) |
| Rural (rural=1, urban=0) | 1.121 | 1.002 |
| | (0.366 - 3.438) | (0.320 - 3.138) |
| Incomplete primary (reference group is no education) | 7.213** | 7.217** |
| | (1.349 - 38.552) | (1.342 - 38.815) |
| Complete primary | 2.080 | 2.004 |
| | (0.237 - 18.272) | (0.226 - 17.763) |
| Incomplete secondary | 2.623 | 2.527 |
| | (0.507 - 13.562) | (0.487 - 13.122) |
| Complete secondary | 1.081 | 1.001 |
| | (0.143 - 8.149) | (0.132 - 7.582) |
| Higher | 0.922 | 0.856 |
| | (0.074 - 11.427) | (0.068 - 10.838) |
| Married/cohabiting (married=1; unmarried=0) | 0.475 | 0.448 |
| | (0.181 - 1.248) | (0.168 - 1.194) |
| Lozi | 0.550 | 0.978 |
| | (0.184 - 1.648) | (0.210 - 4.559) |
| Oshiwambo | 0.169*** | 0.149*** |
| | (0.050 - 0.572) | (0.039 - 0.562) |
| Drank alcohol in the last two weeks (dummy drank=1) | 2.027* | 2.135* |
| | (0.879 - 4.673) | (0.912 - 5.002) |
| Protestant (reference group is Catholic) | 1.544 | 1.541 |
| | (0.598 - 3.989) | (0.583 - 4.071) |
| Other or no religion | 1.533 | 1.278 |
| | (0.430 - 5.462) | (0.340 - 4.807) |
| Within 25km of Angolan border (within 25km =1; >25km=0) | | 0.891 |
| | | (0.322 - 2.464) |
| Within 25km of Zambian border (within 25km =1; >25km=0) | | 0.380 |
| | | (0.109 - 1.324) |
| Constant | 0.003** | 0.002** |
| | (0.000 - 0.290) | (0.000 - 0.285) |
| | 0.013 | 0.014 |

**Table 13: Smoking intensity**

| Independent variables | (1) | (2) |
|---|---|---|
| Observations | 224 | 224 |
| Pseudo R-squared | 0.146 | 0.157 |

Confidence intervals in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Source: Namibian Demographic Health Survey, 2013

# 5. Discussion

## 5.1 Implications of the findings

The results presented in the previous section try to estimate whether the proximity to the Angolan or Zambian borders has had an effect on smoking prevalence or intensity of Wambo, Lozi and Kwangali Namibian men who live within 150km of either border. Logistic regressions are used to test whether, when other factors which affect the likelihood or intensity of smoking are controlled for, proximity to countries which have cheaper cigarettes has an effect. Given that Namibia's northern borders are very porous and cigarettes are significantly cheaper in Angola and Zambia than they are in Namibia then, according to some literature, easy access to cheaper cigarettes should be associated with a higher prevalence of smoking among people who live close to the border compared to those who live further away.

Since the NDHS does not collect data on cigarette prices or income, price elasticities could not be calculated to measure how sensitive consumer demand is to access to cheaper cigarettes. The NDHS data does however collect information on whether one smokes or not, smoking intensity and GPS coordinates. These GPS coordinates were used to create a cut-off point of 150km from either of the northern borders, so that the sample of men used in this analysis were only those who live sufficiently close to these borders. The GPS coordinates also made it possible to calculate the distance of each survey respondent from the borders, albeit with some error due to displacement. Logistic regressions were used to determine whether, once all other factors which are known to influence smoking according to the literature are accounted for, the odds of smoking would be affected by the proximity to the northern border. Since there is a significant difference in smoking prevalence between Lozi, Kwangali and Wambo men, logistic regressions were also run on each group separately to see whether the effects of proximity differed by ethnic group.

The results from tables 8 to 13 show that, of the variables which have been found to be determinants of smoking in other studies, it is only age and alcohol which seem to have a consistently significant effect across ethnic groups. Age has a strong quadratic relationship with the likelihood of smoking, and reporting that one drank alcohol within two weeks of being surveyed has a strong, positive effect on the likelihood of being a smoker. The results for the SES variables, employment, wealth, education, and location, are mixed, which is to be expected given the findings in the literature. Among Kwangali men the SES variables show little effect on the likelihood of smoking. Although some odds ratios for wealth index and education are significant at the 10% level. Among Lozi men being employed is associated with an increased likelihood of smoking at a statistically significant level, and those in the highest two wealth quintiles are significantly less likely to smoke than those in the poorest. Education and location (urban/rural) have no statistically significant effect on smoking. Among Wambo men, employment is also significantly positively associated with the likelihood of smoking, while those with higher levels of education are associated with a lower probability of smoking than those with no education. Wealth and location have no significant effect on the likelihood of smoking, for Wambo men. The results also show that there is no statistically significant relationship between the proximity of the men in the sample to the northern borders and their odds of smoking.

Since no relationship was found between proximity and smoking *prevalence*, another logistic regression was run to test if smoking *intensity* was influenced by proximity to the northern borders. The rationale was that even if the prevalence of smoking was not affected, it may be that those who smoke already would smoke more if they had access to cheaper cigarettes. Although alcohol is found to have a strong positive relationship with smoking intensity, few other odds ratios have meaningful interpretation and again proximity was found to have no relationship with smoking intensity. LR tests were performed on all the logistic regressions and they showed that the addition of the distance variables to the regressions does not have a statistically significant effect on the explanatory power of the model.

Given the findings in the literature, these results are unexpected. There is much evidence from studies in the US, Canada and the EU which show that porous borders combined with cheap tobacco products in neighbouring countries, or neighbouring states in the case of the US, increases the odds of cross border purchase and/or “bootlegging” (Baltagi & Goel, 1987; Sweanor & Martial, 1994; Chaloupka & Warner, 2000; Agaku et al 2015). Since it is well established in tobacco economics literature that increasing the real price of cigarettes leads to a reduction in consumption (Chaloupka & Warner, 2000; Chaloupka, Yurekli, & Fong, 2012), one would expect that those who have easier access to cheaper cigarettes, as a result of proximity, would smoke more than those who do not. However, this is not the case in this study.

This result may be due to a number of things. It could be that proximity to the northern border truly has no effect on smoking prevalence and intensity at all. This is encouraging, because it may mean that the Namibian government can increase taxes much more than it already has without cross border purchases and smuggling being a problem. This is an important result considering that increasing tobacco tax is considered the most effective method of reducing tobacco consumption, and tobacco companies often discourage it on the grounds that higher taxes create incentives for illicit trade in tobacco products (Blecher, 2010).

It is also important to note that although the price paid to the seller is a major determinant in whether people choose to engage in cross border purchases or buying illegal cigarettes, one also needs to consider the "full price", as discussed by Ross (2015). It may be that smokers are not purchasing cheaper cigarettes in Angola or Zambia for personal use, or buying ones which are smuggled into Namibia, because the "full price" of these cigarettes is higher than the normal retail price. This is because the transactions costs could be greater than the price difference between the low tax or illegal cigarettes and full tax/legal cigarettes. This makes the full price of the illegal or low tax cigarettes higher than that of the legal or full-tax cigarettes (Ross, 2015).

It may also be that the price difference may be causing a substitution effect which is not visible. It could be that access to cheaper cigarettes is not leading to an increase in the number of smokers, but existing smokers are simply switching to cheaper cigarettes from across the border. Meaning that the prevalence of smoking and smoking intensity may not have changed, but the substitution effect may have still happened. This effect would not have been picked up in this survey as it does not ask questions of expenditure on cigarettes, brands smoked and where they are purchased from.

The nature of illicit trade, is that it often follows trade routes. Just because cigarettes are cheaper in Angola and Zambia, does not mean that that is where cigarettes will be smuggled from. It may be that, like in South Africa, those who smuggle illegal, cheaper cigarettes may specialise in trade routes and not in a particular product per se (Blecher, 2010). For example, if smugglers specialise in routes between Namibia and the Middle East, it may be that cigarettes from that region will be smuggled into the country simply because of the route, not necessarily the proximity to cheap cigarettes.

## 5.2 Study limitations

As with any research there are limitations to how much can be deduces from these results. In this study this is mainly due to three things, the finite sample properties of maximum likelihood estimators (MLE), the nature of the NDHS dataset, and measurement error.

### 5.2.1 Sample size

Although MLE has many desirable asymptotic properties in terms of efficiency and unbiasedness, over smaller samples they are prone to type II errors. A type II error is a false negative result. It means that one has failed to reject the null hypothesis when it is in fact false (Hart & Clark, 1999). Although our overall sample is about 1676 men, when we separate out the regressions by ethnic groups, the number of observations per regression become much smaller. When separate regressions are run for of Lozi and Kwangali men, who also happen to have the highest smoking prevalence, there are only 276 and 251 observations per regression, respectively. Although this may seem large enough, for MLE estimators the sample size required to avoid type II errors increases, as the number of independent variables in the model increases. Hart and Clark (1999) ran Monte Carlo simulations, on data with a known relationship, to determine how large a sample size was required to prevent type II errors, subject to the number of independent variables, at the 5% significance level. They found that at the 5% significance level, when one has three independent variables in the maximum likelihood regression model, one needs at least 130 observations to have fewer than 5% type II errors. The sample size requirement increases to 180 to 190 observations when there are five independent variables in the model. The models estimated in in this study have twelve independent variables therefore, even though a sample size greater than 200 observations seams large for such a regression, they may not be large enough to avoid false negative findings. (Hart & Clark, 1999).

The main driver of this result, is that smaller sample sizes have larger standard errors, resulting in a ratio of the parameter estimate ($\beta_i$) and the standard error being smaller than the critical value at the 5% level of significance (1.96). This results in a failure to reject the null hypothesis, which in this case is that, proximity to either of the northern borders has no effect on smoking prevalence among men who live within 150km of these borders. (Hart & Clark, 1999).

Hart and Clark (1999) also conclude that one should have between 30 and 50 observations per variable to prevent type two errors. Since there are only 276 for Lozi men and 251 for Kwangali men, there are only about 23 and 21 observations per variable respectively, given that there are twelve variables in the model. This is lower than the recommended number. This highlights a problem not often discussed in social science research, where MLE are used and a large number of variables are included to model relationships as accurately as possible. (Hart & Clark, 1999).

Fortunately sample size is not a problem in the regression which include men from all three ethnic groups, or the regression for Wambo men only. Smaller sample sizes also do not have much of an impact on the likelihood of making type I errors, so it is unlikely that the statistically significant results were misleading even in the regressions with fewer observations (Hart & Clark, 1999).

### 5.2.2 Dataset

The NDHS is not a survey specifically designed for analysing tobacco use. This means that the NDHS survey may not be a sensitive enough instrument to measure whether respondents engage in legal cross border purchases or purchase smuggled cigarettes. In addition to this, men are under sampled in NDHS, probably because the main purpose of this survey is to collect information on, “fertility, family planning, and maternal and child health” (MoHSS and ICF International, 2014). A larger sample of men, particularly Lozi and Kwangali men, would be useful for more accurately testing the hypothesis, as mentioned above. There are only six questions in the adult survey for men which address tobacco use (MoHSS and ICF International, 2014). None of these questions relate to how much of one’s income is spent on cigarettes, where cigarettes are purchased, the price paid for cigarettes or brands smoked. This is all information which would be useful in better understanding cross border purchase or smuggling behaviour.

A more general limitation of survey data is that it relies on self-reported data, which is known to have the potential to be biased particularly when respondents are asked to recall how many cigarettes they smoke or how much of other tobacco products they use (Bilano, et al., 2015).

### 5.2.3 Measurement error

As discussed in the methodology, displacement of GPS coordinates to protect the identity of respondents creates measurement error in the variables of interest, the distance variables. Since GPS coordinates can be displaced up to 10km from its true location, this means that the distance categories created to measure distance from the border, may not reflect reality. Assuming that the positional error is normally distributed, this measurement error is estimated to affect about 4.6% of the sample[7].

Despite these limitations, the best efforts were made to investigate the research question as rigorously as possible given what the NDHS 2013 dataset could offer. Creative methods were employed to convert the GPS coordinates data into distance variables in order to get some measure of access to cheap cigarettes[8].

## 5.3 Policy recommendations

Although the result do not show a relationship between proximity to the northern borders and smoking prevalence mong men, much can still be learned from these results. The quality of data available is important and the absence of it, as can be seen in this paper, may limit one's ability to obtain more detailed and accurate answers to important research questions related to tobacco control. It would be useful to invest in conducting surveys, such as the ITC Surveys, which are conducted periodically and have the specific purpose of investigating tobacco use. These surveys are designed to assess the "psychosocial and behavioural factors" affecting tobacco users (ITC Project, 2014a), and they are particularly useful for assessing whether anti-smoking legislation and public health interventions work. For instance one would be able to ask specific questions around the effectiveness of pictorial health warnings, the effects of taxation and pricing on consumption, perceptions of anti-smoking laws, advertising and other factors which affect smoking or tobacco use behaviour. This will allow for not only more nuanced analysis, but also better targeted interventions through evidence based research.

Surveys which are designed specifically for assessing smoking behaviour are very useful, but cost may be a hindrance to investing in such a study. An almost costless alternative for would be to include questions on expenditure on tobacco products in Namibia's already existing Household Income and Expenditure Survey (IES). At the moment there are no questions around expenditure on tobacco products or alcohol in the IES (Central Bureau of Statistics: National Planning Commision, 2010). This information can be used to measure the sensitivity of demand for cigarettes and other tobacco products to price and income changes. With this data, one can also see whether those who live in places near the northern border are less

[7] Refer to appendix C.
[8] Refer to appendix D

sensitive to price changes, because they have access to cheaper cigarettes from Angola or Zambia.

### 5.4 Potential areas of future research

Possibly due to the lack of data, there has been little research done on tobacco control in Namibia. The pictorial warnings which came into effect in 2015 provide an excellent opportunity to look at how these warning affect the smoking behaviour of Namibian smokers, and the prevalence of cross border purchases of cigarettes and illicit trade. Pack collection in areas along the northern Namibian border would make it possible to easily determine whether the cigarettes smoked in those areas originated from Namibia, as none of Namibia's northern neighbours require pictorial warnings on tobacco products. Qualitative research could also be done on the perceptions of these warnings around the northern border to assess whether they encourage smoking cessation, prevent uptake or simply encourages the purchase of foreign tobacco products. It would also be interesting to compare the tobacco use prevalence rates using the 2013 NDHS and future NDHS studies, to see if the pictorial warnings have had a significant effect on smoking prevalence or intensity.

This study has shown that there are significant differences in smoking rates by ethnic group. It would be worthwhile to understand what cultural factors may influence tobacco use. Due to the nuanced nature of this type of question, a more qualitative approach may be required. Information from such a study may be useful for tailoring public health interventions around the needs and beliefs of communities and it may provide information which can better equip health workers to encourage smoking cessation or prevent uptake.

# 6. Conclusion

Since Namibia's ratification of the WHO FCTC, considerable progress has been made in passing regulation protecting the public from the global health epidemic, despite intimidation through threat of legal action by tobacco companies (Framework Convention Alliance, 2012; Tam & van Walbeek, 2013; Tavernise, 2013). The implementation of mandatory pictorial warnings on all tobacco products came into effect on the 1st of April 2015 in Namibia, which shows the government's commitment to the WHO FCTC (World Health Organisation Regional Office for Africa, 2015).

Although this paper does not find a significant relationship between proximity to the northern borders and smoking prevalence, this result is still important. This result may support the finding of other studies that high taxes on tobacco products, such as cigarettes, are not necessarily related to an increase in illicit trade in tobacco, as tobacco companies often like to argue. These findings suggest that despite the fact that cigarettes prices are higher in Namibia than they are in Angola and Zambia, due to higher taxes, and that the northern border is very porous, this has not had an effect on the odds or intensity of smoking. Therefore the Namibian government can go even further with its tobacco control policy by increase taxes on tobacco products, without fear of encouraging illicit trade or cross border purchase.

# 7. References


MoHSS and ICF International, 2014. *Namibia Demographic Health Survey 2013 [Dataset].* Windhoek, Namibia, and Rockville, Maryland, USA: MoHSS and ICF International.

Agaku, I. et al., 2015. Impact of cigarette price differences across the entire European Union on cross-border purchase of tobacco products among adult cigarette smokers. *Tobacco Control,* pp. 1-8.

Baltagi, B. & Goel, R., 1987. Quasi-Experimental Price Elasticities of Cigarette Demand and the Bootlegging Effect. *American Journal of Agricultural Economics,* 69(4), pp. 750-754.

Baltagi, B. & Levin, D., 1986. Estimating Dynamic Demand for Cigarettes Using Panel Data: The Effects of Bootlegging, Taxation and Advertising Reconsidered. *The Review of Economics and Statistics,* 68(1), pp. 148-155.

Bilano, V. et al., 2015. Global trends and projections for tobacco use, 1990–2025: an analysis of smoking indicators from the WHO Comprehensive Information Systems for Tobacco Control. *Lancet,* 385(9972), pp. 966-976.

Blecher, E., 2010. A mountain or a molehill: is the illicit trade in cigarettes undermining tobacco control policy in South Africa?. *Trends in Organised Crime,* Volume 13, p. 299–315.

Blecher, E. & Drope, J., 2014. The rewards, risks and challenges of regional tobacco tax harmonisation. *Tobacco Control,* pp. 1-5.

Blecher, E. & van Walbeek, C., 2004. An international analysis of cigarette affordability. *Tobacco Control,* Volume 13, p. 339–346.

Brophy, T., Daniels, R. & Musundwa, S., 2014. *gpsbound: Routine for importing and verifying geographical information from a user provided shapefile,* Cape Town: SALDRU, University of Cape Town: A Southern Africa Labour and Development Research Unit Working Paper Number 132.

Central Bureau of Statistics: National Planning Commision, 2010. *Namibia Household Income and Expenditure Survey 2009/10,* Windhoek: National Planning Commision.

Centres for Diesease Control and Prevention, 1998. Response to Increases in Cigarette Prices by Race/Ethnicity, Income, and Age Groups — United States, 1976–1993. *Morbidity and Mortality Weekly Report*, 31 July, 47(29), pp. 605-609.

Chaloupka, F. & Warner, K., 2000. The economics of smoking. *Handbook of health economics,* Volume 1, pp. 1539-1627.

Chaloupka, F., Yurekli, A. & Fong, G., 2012. Tobacco taxes as a tobacco control strategy. *Tobacco Control,* Volume 21, pp. 172-180.

Chelwa, G., van Walbeek, C. & Blecher, E., 2015. *Evaluating South Africa's Tobacco Control Initative: A Synthetic Control Approach.* s.l.:ERSA working paper 567 .

Customs and Excise Act 20 , 1998. s.l.:s.n.

DIVA-GIS, n.d.. *Download data by country.* [Online]
Available at: http://www.diva-gis.org/gdata
[Accessed 20 September 2015].

ESRI, 2008. *ESRI Shapefile Technical Description: An ESRI White Paper,* Redlands, CA : Environmental Systems Research Institute.

Ezeamalu, B., 2015. *How American company, Philip Morris, is flooding Nigeria with 122 million 'units' of cigarettes.* [Online]
Available at: http://www.premiumtimesng.com/business/194867-how-american-company-philip-morris-is-flooding-nigeria-with-122-million-units-of-cigarettes.html

Framework Convention Alliance, 2012. *BAT threatens Namibia's tobacco control law.* [Online]
Available at:
http://www.fctc.org/images/stories/WNTD_CaseStudy_Layout_Namibia_EN.pdf
[Accessed 12 June 2015].

Gallet, A., 2015. *How Cigarette Smuggling Funds al-Qaeda in Africa.* [Online]
Available at: http://africatimes.com/2015/12/04/how-cigarette-smuggling-funds-al-qaeda-in-africa/

Gbubemi, G., 2015. *Tobacco smuggling costs Nigeria over N250 billion in unpaid taxes annually – Freddy Messanvi.* [Online]
Available at: http://dailytimes.com.ng/tobacco-smuggling-costs-nigeria-over-n250-billion-in-unpaid-taxes-annually-freddy-messanvi/

Griffiths, I., 1986. The Scramble for Africa: Inherited Political Boundaries. *The Geographical Journal,* 152(2), pp. 204-216.

Gueye, C. et al., 2014. Namibias path towards malaria elimination: a case study of malaria strategies and costs along the northern border. *BioMed Central,* Volume 14, pp. 1-16.

Guindon, G. & Boisclair, D., 2003. Past, current and future trends in tobacco use. *HNP Discussion Paper: Economics of Tobacco control No. 6.*

Hart, R. & Clark, D., 1999. Does size matter? Exploring the small sample properties of maximum likelihood estimators.

Hosmer, D. & Lemenshaw, S., 2000. *Applied Logistic Regression: Second Edition.* New York, Toronto: John Whiley & Sons, Inc.

Husten, C., 2009. How should we define light or intermittent smoking? Does it matter?. *Nicotine & Tobacco Research,* p. ntp010.

ITC Project, 2014a. *About ITC.* [Online]
Available at: http://www.itcproject.org/about
[Accessed 24 January 2016].

ITC Project, 2014b. *ITC Zambia National Report. Findings from the Wave 1 (2012) Survey,* Waterloo, Ontario Canada: University of Waterloo.

ITC Project, 2015. *Zambia: Timeline of Tobacco Control Policies and ITC Surveys.* [Online]
Available at: http://www.itcproject.org/galleria/ZM(MB-Sept4,2015).pdf

Jha, P., Chaloupka, F., Corrao, M. & Jacob, B., 2006. Reducing the burden of smoking world-wide: effectiveness of interventions and their coverage. *Drug and Alcohol Review,* Volume 25, p. 597–609.

Jha, P. & Peto, R., 2014. Global Effects of Smoking, of Quitting, and of Taxing Tobacco. *The New England Journal of Medicine,* 370(1), pp. 60-68.

John, R., Mamudu, H. & Liber, A., 2012. Socioeconomic Implications of Tobacco Use in Ghana. *Nicotine & Tobacco Research,* 14(10), p. 1205–1212.

King, G. & Zeng, L., 2001. Logistic Regression in Rare Events Data. *Political analysis,* 9(2), pp. 137-163.

Koyana, X., 2015. *SA economy losing billions due to illegal ciagarette trade.* [Online]
Available at: http://ewn.co.za/2015/04/24/SA-economy-losing-billions-due-to-illegal-cigarettes-trade

Lovenheim, M., 2008. How Far to the Border?: The Extent and Impact of Cross-Border Casual Cigarette Smuggling. *National Tax Journal,* 61(1), pp. 7-33.

Lusaka Times, 2015. *ZRA unearths scam of smuggled cigarettes worth K6 million at Katima Mulilo border.* [Online]
Available at: https://www.lusakatimes.com/2015/07/20/zra-unearths-scam-of-smuggled-cigarettes-worth-k6-million-at-katima-mulilo-border/

Lusaka Voice, 2015. *Govt. to tighten security on cigarette smuggling.* [Online]
Available at: http://lusakavoice.com/2015/04/14/govt-to-tighten-security-on-cigarette-smuggling-2/

Maps of the World, 2015. *Political Map of Namibia.* [Online]
Available at: http://www.mapsofworld.com/namibia/namibia-political-map.html
[Accessed 13 February 2016].

MEASURE DHS/ICF International, 2013. *Standard Recode Manual for DHS 6,* Rockville, Maryland, USA: ICF International.

Merriman, D., 2012. *Understand, measure and combat Tobacco Smuggling,* Washington: World Bank Economics of Tobacco Toolkit, 7.

Moonga, B., 2011. *Current politics of tobacco control in Zambia.* [Online]
Available at: http://www.zambiawatchdog.com/the-politics-of-tobacco-control-in-zambia/
[Accessed 22 January 2016].

Namibia Statistics Agency (NSA), 2014. *Namibia Population Projections 2011-2041,* Windhoek: NSA.

Namibia Trade Information Portal, 2015. *Traveller's allowances.* [Online]
Available at: https://www.namibiatradeinformationportal.gov.na/traders-toolkit/travellers-allowances.php

Namibian Sun, 2013. *Namibia, Angola border cards coming.* [Online]
Available at: http://www.namibiansun.com/business/namibia-angola-border-cards-coming.53774
[Accessed 25 November 2015].

New Era, 2007. *'Human Taxis' Rife at Northern Borders.* [Online]
Available at: 'Human Taxis' Rife at Northern Borders
[Accessed 21 January 2016].

New Era, 2015. *Two cigarette smugglers in court.* [Online]
Available at: https://www.newera.com.na/2015/05/12/cigarette-smugglers-court/

News Ghana, 2016. *Zambia refuses to ban tobacco production.* [Online]
Available at: http://newsghana.com.gh/zambia-refuses-to-ban-tobacco-production/
[Accessed 22 January 2016].

Pampel, F., 2005. Patterns of Tobacco Use in the Early Epidemic Stages: Malawi and Zambia, 2000–2002. *American Journal of Public Health,* 95(6), pp. 1009-1015.

Pampel, F., 2008. Tobacco use in Sub-Saharan Africa: Estimates from demographic health surveys. *Social Science & Medicine,* Volume 66, pp. 1772-1783.

Pampel, F., 2008. Tobacco use in Sub-Saharan Africa: Estimates from demographic health surveys. *Social Science and Medicine,* Volume 66, pp. 1772-1783.

Peer, N., Bradshaw, D., Laubscher, R. & Steyn, K., 2009. Trends in adult tobacco use from two South African demographic and health surveys conducted in 1998 and 2003. *South African Medical Journal,* Volume 99, pp. 744-749.

Ross, H., 2015. *Understnding and measuring cigarette tax avoidance an evasion: A methodological guide,* s.l.: Tobacconomics.

SACU, 2013. *About SACU: History of SACU.* [Online]
Available at: http://www.sacu.int/show.php?id=394

Saloojee, Y., 1995. Price and income elasticity of demand for cigarettes in South Africa. In: *Tobacco and Health.* s.l.:Springer US, pp. 235-239.

Sasman, C., 2013. *Noose to tighten on contraband cigarettes.* [Online]
Available at: http://www.namibian.com.na/index.php?id=103974&page=archive-read

Shea, O. & Guillermo, R., 2006. *Guide to DHS Statistics,* Calverton, Maryland : Demographic and Health Surveys, ORC Macro .

Simon, D., 1998. Desert enclave to regional gateway? The reintegration of Walvis Bay into Namibia. In: *South Africa in Southern Africa: reconfiguring the region.* s.l.:s.n., pp. 103-128.

Stockburger, D., 1996. *Chi-Square and tests of contingency tables.* [Online]
Available at: http://www.psychstat.missouristate.edu/introbook/sbk28m.htm
[Accessed 4 January 2016].

Stoklosa, M. & Ross, H., 2013. Contrasting academic and tobacco industry estimates of illicit cigarette trade: evidence from Warsaw, Poland. *Tobacco control.*

Sweanor, D. & Martial, L., 1994. The smuggling of tobacco products: lessons from Canada. *Non-Smokers Rights Association.*

Tam, J. & van Walbeek, C., 2013. Tobacco control in Namibia: the importance of government capacity, media coverage and industry interference. *Tobacco control,* p. 1–6.

Tavernise, S., 2013. *Tobacco Firms' Strategy Limits Poorer Nations' Smoking Laws.* [Online]
Available at: http://www.njgasp.org/wp-content/uploads/2014/06/NYTimes_Tobacco_Firms_Strategy_Limits_Poor_nations_Smoking_Laws_12-13-13.pdf
[Accessed 12 June 2015].

The DHS Program, 2013. *GPS Data Collection.* [Online]
Available at: https://dhsprogram.com/What-We-Do/GPS-Data-Collection.cfm
[Accessed 20 September 2015].

The Nambia Ministry of Health and Social Services (MoHSS) and ICF International, 2014. *The Namibia Demographic and Health Survey 2013,* Windhoek, Namibia, and Rockville, Maryland, USA: MoHSS and ICF International.

The Tobacco Atlas, 2015. *Angola.* [Online]
Available at: http://www.tobaccoatlas.org/country-data/angola/

Townsend, J., Roderick, P. & Cooper, J., 1994. Cigarette Smoking By Socioeconomic Group, Sex, And Age: Effects Of Price, Income, And Health Publicity. *British Medical Journal,* 309(6959), pp. 923-927.

Townsend, L., Fisher, A., Gilreath, T. & King, G., 2006. A systematic literature review of tobacco use among adults 15 years and older in Sub-Saharan Africa. *Drug and Alcohol Dependence,* Volume 84, pp. 14-27.

Tumwine, J., 2011. Implementation of the Framework Convention on Tobacco Control in Africa: Current Status of Legislation. *International Journal of Environmental Research and Public Health,* Volume 8, pp. 4312-4331.

UCLA: Statistical Consulting Group, n.d. a. *Stata Annotated Output: Ordered Logistic Regression.* [Online]
Available at: http://www.ats.ucla.edu/stat/stata/output/stata_ologit_output.htm
[Accessed 2 January 2016].

UCLA: Statistical Consulting Group, n.d. b. *Stata Data Analysis Examples: Ordered Logistic Regression.* [Online]
Available at: http://www.ats.ucla.edu/stat/stata/dae/ologit.htm
[Accessed 3 January 2016].

UCLA: Statistical Consulting Group, n.d. *Stata Data Analysis Examples: Ordered Logistic Regression.* [Online]
Available at: http://www.ats.ucla.edu/stat/stata/dae/ologit.htm
[Accessed 12 January 2016].

Umraw, A., 2015. *PMB a hub for cig smuggling.* [Online]
Available at: http://www.news24.com/SouthAfrica/News/PMB-a-hub-for-cig-smuggling-20150704

United Nations, 2016. *Chapter IX Health: WHO Framework Convention on Tobacco Control.* [Online]
Available at: https://treaties.un.org/pages/ViewDetails.aspx?src=TREATY&mtdsg_no=IX-4&chapter=9&lang=en
[Accessed 20 January 2016].

van Walbeek, C., 2002. Recent trends in smoking prevelance in South Africa - some evidence from AMPS data. *South African Medical Journal,* 92(6), pp. 468-472.

van Walbeek, C., 2005. *The economics of tobacco control in South Africa.* s.l.:University of Cape Town .

Williams, R., 2004. *Normal distribution.* [Online]
Available at: https://www3.nd.edu/~rwilliam/stats1/x21.pdf
[Accessed 7 February 2016].

Wooldridge, J., 2013. *Introductory Econometrics: A modern approach.* 5th edition ed. Canada: South-Western Cengage Learning.

World Health Organisation Regional Office for Africa, 2015. *Namibia: Pictorial health warnings on tobacco products from 1st April 2015.* [Online]
Available at: http://www.afro.who.int/en/clusters-a-programmes/hpr/health-risk-factors/tobacco/highlights/4603-namibia-pictorial-health-warnings-on-tobacco-products-from-1st-april-2015.html
[Accessed 20 January 2016].

World Health Organisation, 2011. *WHO Report on the Global Tobacco Epidemic, 2011: Warnigs about the dangers of tobacco,* Geneva: WHO Press.

World Health Organisation, 2015a. *WHO Report on the Global Tobacco Epidemic, 2015: Raising taxes on tobacco,* Geneva: WHO Press.

World Health Organisation, 2015b. *The WHO Framework Convention on Tobacco Control: an overview.* [Online]
Available at: http://www.who.int/fctc/Protocol_summary_en.pdf?ua=1
[Accessed 20 Jarnuary 2016].

World Health Organisation, 2015c. *WHO Report on the Global Tobacco Epidemic, 2015: country profile Angola.* [Online]
Available at: http://www.who.int/tobacco/surveillance/policy/country_profile/ago.pdf
[Accessed 12 July 2015].

World Health Organisation, 2015d. *WHO Report on the Global Tobacco Epidemic, 2015: country profile Zambia.* [Online]
Available at: http://www.who.int/tobacco/surveillance/policy/country_profile/zmb.pdf

Zeller, W., 2010. Neither arbitrary nor artificial: Chiefs and the making of the Namibia-Zambia borderland. *Journal of Borderlands Studies,* 25(2), pp. 6-21.

# Appendix A: Contingency tables[9]

| Table 1: Smoker by Employment | | | | |
|---|---|---|---|---|
| | | **Employment** | | |
| | | 0. Unemployed | 1. Employed | Total |
| **Smoker** | 0. non-smoker | 811 | 632 | 1443 |
| | % | 91.64 | 80.3 | 86.3 |
| | 1. Smoker | 74 | 155 | 229 |
| | % | 8.36 | 19.7 | 13.7 |
| | Total | 885 | 787 | 1672 |
| | % | 100 | 100 | 100 |
| | Pearson chi2(1) | 45.2666 | Pr = 0.001 | |

| Table 2: Smoker by wealth index | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | **Wealth index** | | | | | |
| | | 1. poorest | 2. poorer | 3. middle | 4. richer | 5. richest | Total |
| **Smoker** | 0. non-smoker | 417 | 366 | 359 | 228 | 76 | 1446 |
| | % | 83.9 | 85.2 | 88.21 | 90.48 | 85.39 | 86.33 |
| | 1. Smoker | 80 | 64 | 48 | 24 | 13 | 229 |
| | % | 16.1 | 14.88 | 11.79 | 9.52 | 14.61 | 13.67 |
| | Total | 497 | 430 | 407 | 252 | 89 | 1675 |
| | % | 100 | 100 | 100 | 100 | 100 | 100 |
| | Pearson chi2(4) | 7.967 | Pr = 0.093 | | | | |

| Table 3: Smoker by location | | | | |
|---|---|---|---|---|
| | | **Rural** | | |
| | | 0. Urban | 1. Rural | Total |
| **Smoker** | 0. non-smoker | 294 | 1152 | 1446 |
| | % | 83.76 | 87.01 | 86.33 |
| | 1. Smoker | 57 | 172 | 229 |
| | % | 16.24 | 12.99 | 13.67 |
| | Total | 351 | 1324 | 1675 |
| | % | 100 | 100 | 100 |
| | Pearson chi2(1) | 2.4805 | Pr = 0.115 | |

[9] The source of the data for all contingency tables is the NDHS 2013 dataset.

| Table 4: Smoker by education | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | **Education** | | | | | | |
| | | 1. no education | 2. incomplete primary | 3. complete primary | 4. incomplete secondary | 5. complete secondary | 6. higher | Total |
| **Smoker** | 0. non-smoker | 132 | 365 | 115 | 608 | 144 | 82 | 1446 |
| | % | 81.99 | 87.11 | 86.47 | 87.23 | 82.76 | 90.11 | 86.33 |
| | 1. Smoker | 29 | 54 | 18 | 89 | 30 | 9 | 229 |
| | % | 18.01 | 12.89 | 13.53 | 12.77 | 17.24 | 9.89 | 13.67 |
| | Total | 161 | 419 | 133 | 697 | 174 | 91 | 1675 |
| | % | 100 | 100 | 100 | 100 | 100 | 100 | 100 |
| | | Pearson chi2(5) | 6.2529 | Pr = 0.282 | | | | |

| Table 5: Smoker by marital status | | | | |
|---|---|---|---|---|
| | | **Married** | | |
| | | 0. unmarried | 1. married | Total |
| **Smoker** | 0. non-smoker | 1071 | 375 | 1446 |
| | % | 87.86 | 82.24 | 86.33 |
| | 1. Smoker | 148 | 81 | 229 |
| | % | 12.14 | 17.76 | 13.67 |
| | Total | 1219 | 456 | 1675 |
| | % | 100 | 100 | 100 |
| | Pearson chi2(1) | 8.8873 | Pr = 0.003 | |

| Table 6: Smoker by drinking alcohol | | | | |
|---|---|---|---|---|
| | | **Drinks alcohol** | | |
| | | 0. non-drinker | 1. drinker | Total |
| **Smoker** | 0. non-smoker | 903 | 543 | 1446 |
| | % | 90.75 | 79.85 | 86.33 |
| | 1. Smoker | 92 | 137 | 229 |
| | % | 9.25 | 20.15 | 13.67 |
| | Total | 995 | 680 | 1675 |
| | % | 100 | 100 | 100 |
| | Pearson chi2(1) | 40.6689 | Pr = 0.000 | |

**Table 7: Smoker by respondent language**

| | | Language | | | |
|---|---|---|---|---|---|
| | | 1. Kwangali | 2. Lozi | 3. Oshiwambo | Total |
| **Smoker** | 0. non-smoker | 204 | 185 | 1057 | 1446 |
| | % | 80.63 | 66.79 | 92.31 | 86.33 |
| | 1. Smoker | 49 | 92 | 88 | 229 |
| | % | 19.37 | 33.21 | 7.69 | 13.67 |
| | Total | 253 | 277 | 1145 | 1675 |
| | % | 100 | 100 | 100 | 100 |
| | Pearson chi2(2) | 131.3394 | Pr = 0.000 | | |

**Table 8: Smoker by religion**

| | | Religion | | | |
|---|---|---|---|---|---|
| | | 1. Roman Catholic | 2. Protestant | 3. Other/no religion | Total |
| **Smoker** | 0. non-smoker | 368 | 980 | 97 | 1446 |
| | % | 83.45 | 88.85 | 74.62 | 86.33 |
| | 1. Smoker | 73 | 123 | 33 | 229 |
| | % | 16.55 | 11.15 | 25.38 | 13.67 |
| | Total | 441 | 1103 | 130 | 1675 |
| | % | 100 | 100 | 100 | 100 |
| | Pearson chi2(2) | 24.1377 | Pr = 0.000 | | |

**Table 9: Smoker by distance from the Angolan border**

| | | Distance from Angolan border | | | |
|---|---|---|---|---|---|
| | | 1. Within 25km | 2. Between 25km and 50km | 3. More than 50km | Total |
| **Smoker** | 0. non-smoker | 485 | 404 | 557 | 1446 |
| | % | 88.67 | 89.98 | 82.03 | 86.33 |
| | 1. Smoker | 62 | 45 | 122 | 229 |
| | % | 11.33 | 10.02 | 17.97 | 13.67 |
| | Total | 547 | 449 | 679 | 1675 |
| | % | 100 | 100 | 100 | 100 |
| | Pearson chi2(2) | 18.2153 | Pr = 0.000 | | |

| Table 10: Smoker by distance from the Zambian border | | | | | |
|---|---|---|---|---|---|
| | | Distance from Zambian border | | | |
| | | 1. Within 25km | 2. Between 25km and 50km | 3. More than 50km | Total |
| **Smoker** | 0. non-smoker | 142 | 33 | 1271 | 1446 |
| | % | 69.61 | 70.21 | 89.26 | 86.33 |
| | 1. Smoker | 62 | 14 | 153 | 229 |
| | % | 30.39 | 29.79 | 10.74 | 13.67 |
| | Total | 204 | 47 | 1424 | 1675 |
| | % | 100 | 100 | 100 | 100 |
| | Pearson chi2(2) | 69.004 | Pr = 0.000 | | |

# Appendix B: Smokers by language spoken

**Table 1: Smokers by language spoken: Male**

| | Kwangali | | Losi | | Oshiwambo | | Afrikaans | | Damara | | English | | Herero | | Other | | Total | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | n=266.17 | weighted % | n=188.86 | weighted % | n=1263.68 | weighted % | n=3.90 | weighted % | n=4.10 | weighted % | n=28.68 | weighted % | n=18.63 | weighted % | n=63.98 | weighted % | n=1838 | weighted % |
| Non-smoker | 216.31 | 81.27 | 126.82 | 67.15 | 1166.29 | 92.29 | 2.99 | 76.63 | 2.52 | 61.43 | 17.67 | 61.61 | 16 | 83.82 | 50.32 | 78.66 | 1596.46 | 86.86 |
| Smoker | 49.86 | 18.73 | 62.04 | 32.85 | 97.39 | 7.71 | 0.91 | 23.37 | 1.58 | 38.57 | 11.01 | 38.39 | 3 | 16.18 | 13.66 | 21.34 | 241.54 | 13.14 |

Source: NDHS, 2013

**Table 2: Smokers by language spoken: Female**

| | Kwangali | | Losi | | Oshiwambo | | Afrikaans | | Damara | | English | | Herero | | Other | | Total | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | n=623.71 | weighted % | n=372.47 | weighted % | n=2880.13 | weighted % | n=7.50 | weighted % | n=5.04 | weighted % | n=56.85 | weighted % | n=62.90 | weighted % | n=175.41 | weighted % | n=4184 | weighted % |
| Non-smoker | 603.98 | 96.84 | 367.04 | 98.54 | 2864.55 | 99.46 | 6.02 | 80.27 | 5.04 | 100.00 | 49.33 | 86.77 | 61.10 | 97.14 | 169.03 | 96.36 | 4126.10 | 98.62 |
| Smoker | 19.72 | 3.16 | 5.43 | 1.46 | 15.57 | 0.54 | 1.48 | 19.73 | 0.00 | 0.00 | 7.52 | 13.23 | 1.80 | 2.86 | 6.38 | 3.64 | 57.91 | 1.38 |

Source: NDHS, 2013

# Appendix C: Estimates of displacement

**Table 1: Proportion of potentially displaced people by category**

| Zambia | | | | | Angola | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | 1st cut off: 25km | number of observations | formula | proportion of observations potentially displaced | | 1st cut off: 25km | number of observations | formula | proportion of observations potentially displaced |
| A | 0-35km | 770 | C/A | 0.308 | I | 0-35km | 2884 | K/I | 0.379 |
| B | 0-30km | 731 | D/B | 0.133 | J | 0-30km | 2643 | L/J | 0.240 |
| C | 15-35km | 237 | | | K | 15-35km | 1093 | | |
| D | 20-30km | 97 | | | L | 20-30km | 633 | | |
| | 2nd cut off: 50km | | | | | 2nd cut off: 50km | | | |
| E | 25-60km | 201 | G/E | 0.483 | M | 25-60km | 2124 | O/M | 0.639 |
| F | 25-55km | 172 | H/F | 0.134 | N | 25-55km | 1735 | P/N | 0.376 |
| G | 40-60km | 97 | | | O | 40-60km | 1358 | | |
| H | 45-55km | 23 | | | P | 45-55km | 652 | | |

**Table 2: Estimates of displacement between distance categories**

| | Zambia | | | | | Angola | | | |
|---|---|---|---|---|---|---|---|---|---|
| | 1st cut off: 25km | number of observations | % potentially displaced | potential error | | 1st cut off: 25km | number of observations | % potentially displaced | potential error |
| -10km | 15-25km | 198 | 0.01 | 1.980 | -10km | 15-25km | 543 | 0.01 | 5.430 |
| -3σ | 20-25km | 97 | 0.0013 | 0.126 | -3σ | 20-25km | 324 | 0.0013 | 0.421 |
| -2σ | 21.67-25km | 75 | 0.0028 | 0.210 | -2σ | 21.67-25km | 123 | 0.0028 | 0.344 |
| -σ | 23.33-25km | 46 | 0.1587 | 7.300 | -σ | 23.33-25km | 0 | 0.1587 | 0.000 |
| σ | 25-26.67km | 0 | 0.1587 | 0.000 | σ | 25-26.67km | 60 | 0.1587 | 9.522 |
| 2σ | 25-28.33km | 0 | 0.0028 | 0.000 | 2σ | 25-28.33km | 179 | 0.0028 | 0.501 |
| 3σ | 25-30km | 0 | 0.0013 | 0.000 | 3σ | 25-30km | 309 | 0.0013 | 0.402 |
| +10km | 25-35km | 39 | 0.01 | 0.390 | +10km | 25-35km | 550 | 0.01 | 5.500 |
| | 2nd cut off: 50km | | | | | 2nd cut off: 50km | | | |
| -10km | 40-50km | 45 | 0.01 | 0.450 | -10km | 40-50km | 845 | 0.01 | 8.450 |
| -3σ | 45-50km | 0 | 0.0013 | 0.000 | -3σ | 45-50km | 528 | 0.0013 | 0.686 |
| -2σ | 46.67-50km | 0 | 0.0028 | 0.000 | -2σ | 46.67-50km | 377 | 0.0028 | 1.056 |
| -σ | 48.33-50km | 0 | 0.1587 | 0.000 | -σ | 48.33-50km | 181 | 0.1587 | 28.725 |
| σ | 50-51.67km | 0 | 0.1587 | 0.000 | σ | 50-51.67km | 0 | 0.1587 | 0.000 |
| 2σ | 50-53.33km | 23 | 0.0028 | 0.064 | 2σ | 50-53.33km | 75 | 0.0028 | 0.210 |
| 3σ | 50-55km | 23 | 0.0013 | 0.030 | 3σ | 50-55km | 124 | 0.0013 | 0.161 |
| +10km | 50-60km | 52 | 0.01 | 0.520 | +10km | 50-60km | 513 | 0.01 | 5.130 |
| | total error | | | 11.071 | | total error | | | 66.538 |
| | **grand total** | **77.609** | | | | | | | |
| | **total sample** | **1676** | | | | | | | |
| | **% potential error** | **0.046** | | | | | | | |

Due to the positional error created by displacement, it is possible that people who are allocated into certain distance categories may in actual fact not fall into those categories. The distance variable categories are defined as 0-25km (category one), 25km-50km (category 2) and 50-150km (category 3) from either the Angolan or Zambian borders. Table 1 shows the proportion of people at each cut off point of the distance variables who could potentially be miscategorised due to displacement. This table is divided into two, on the left hand side of the table are the calculations for Zambia and on the right are those for Angola. Row D of table 1 shows the number of people who live within 5km of either side of the 25km cut off point for category 1 of the distance variable for the Zambian border. This row represents the number of people who are recorded as living within 25km of the border, but due to displacement may in fact actually fall into category one when they are actually in category two, and vice versa. As a result one cannot know for sure which category they fall into. Row B shows the number of people who live within 25km of the Zambian border and those who live 5km beyond the cut off, therefore those who live 30km from the Zambian border. This row represents all the observations which are actually in category one and those which could have originally be in category one if not for displacement. The ratio of these observations, as shown in the column for "proportion of observations potentially displaced", is 0.133. This means that 13.3% of observations within 30km of the Zambian border could be categories incorrectly due to positional error caused by displacement. A similar logic applies to rows A and C. 1% of observations are displaced by 0-10km, as opposed to 0-5km like the others, therefor row C gives us the number of observations which have a 1% chance of being displaced by 10km. Like row B, row A shows all observations which are in category one, along with those which may actually be in category one, but due to displacement end up in category two, and vice versa. The ratio of rows C and A is 0.308. This means that due to positional error from displacement, we are unsure of whether 30.8% of observations within 35km of the Zambian border reflect the true position of the observations. This same interpretation can be applied to the other cut offs for both the Zambian and Angolan borders in table 1.

Table 2 shows the calculation of the estimated number of people who would be grouped into the wrong distance categories, due to displacement if positional error is assumed to be normally distributed. Like table 1, it is divided into two. It is a characteristic of the normal distribution that about 68.26% of all normally distributed observations fall within one standard deviation ($\sigma$) of either side of the mean, 95.44% fall within two standard deviations, and 99.74% fall within three standard deviations either side of the mean (Williams, 2004). If one assumes that each of the displaced GPS coordinates is the centre of the normal distribution we know that, for those that are those that are displaced by 0-5km, the real (non-displaced) GPS coordinate has essentially a 100% probability of being within 5km of it. Therefore 5km is divided by three, and 1.67 is taken to be one standard deviation. As one can see from the table 2 in the column for the "1st cut off: 25km", those within 23.33-25km are within one standard deviation (-σ) of the 25km distance variable cut off. There are 46 observations in this interval and there is a 15.87%[10] chance that some of them will be more than one standard deviation away from 25km. If we multiply the likelihood of falling outside of the interval with the "% potential displacement" one ends up with 7.3 observations which are likely to fall beyond the 25km cut off. We apply the same calculation to those observations from intervals more than one standard deviation away from the respective cut offs, including the 1% that we know have been displaced by 0-10km. When these numbers are added up, there are a grand total of almost 78 observations which are likely to have been misallocated to distance categories, if positional errors were normally distributed.

It should be noted that in table 2, the number of observations are sometimes recorded as zero. This is because there are some intervals which have not observations. This is not an anomaly, it is simply a result of some areas being less densely populated than others and that not all observations being evenly distributed.

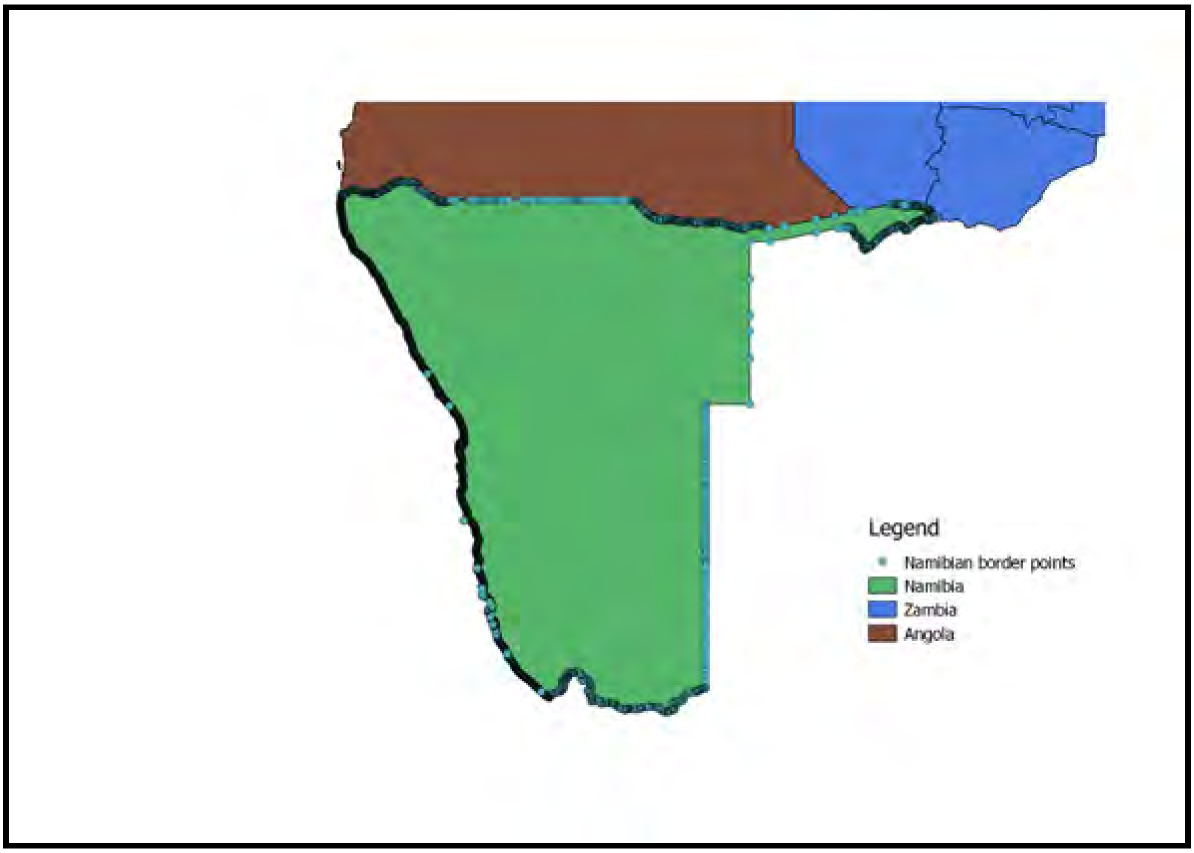
[10] Since there is a 68.26% chance of being within one standard deviation of either side of the mean in a normal distribution, then if one is only looking at one side of the mean, this probability is divided by two.

# Appendix D: Mapping GPS coordinates to create distance variables

Polygon shapefiles describe geographic areas such as border and/or boundaries of administrative areas, such as countries or regions within a country, along with an attribute table. These polygon shapefiles are made up of dots which are joined in clockwise direction. Each polygon shapefile is a collection of files stored in the same directory, which share the same pre-fix. For a file to be a shapefile it must contain at least three file formats, shapefile (*.shp), dBase (*.dbs) file and index (*.shx) file. A shapefile contains the actual GPS coordinates, the dBase file contains a table of attribute features, such as names of cities and town, which are non-spatial, and the index file makes it possible for the GIS software to efficiently navigate the shapefile. These files together store information on geometric location, such as geographically demarcated borders, and its associated attribute information in digital vector format. (ESRI, 2008; Brophy, Daniels, & Musundwa, 2014).

Since the shapefile for the Namibian borders contained the borders for the administrative regions as well, QGIS's "dissolve" option was used to remove the internal borders of the country so that only the outer border remained. Using Stata's "shp2dta" command, the outer borders of the Namibia was traced and converted into GPS coordinates from the shapefile for Namibia. The Namibian border coordinates are then read back into QGIS to create a new shape file of border points which can be seen in Figure 1. Using "toggle editing" in QGIS, the Namibia-Angolan and Namibian-Zambian borders were essentially cut out of the shapefile of Namibian border points, creating new shapefiles for each of these borders. The points for the GPS coordinates of the Angola-Namibia and Zambia Namibia borders are depicted in Figures 2 and 3, respectively.

**Figure 1: Namibian border points**




Again using the Stata's "shp2data" command, the points along the Namibian-Angolan and Namibian-Zambian borders were converted into datasets of GPS coordinates and merges into the larger dataset. Stata's "geonear" command was then used to generate distances between each cluster and the closest point along each border. Figure 4 shows the clusters for the whole country in relation to the Angolan and Zambian borders. The geonear command creates a variable which shows the shortest distance in kilometres between the cluster and the border. Essentially, this command drops a perpendicular line from the cluster to the nearest point on the border of interest and measures the distance of that line. The main limitation of this command is that it does not take into consideration any physical features or geographic barriers between the cluster and the border, such a mountains, bodies of water, or even other countries.

Another limitation is that shp2dta does not always assign GPS coordinates along the borders very close to each other or particularly uniformly when it traces the border as can be seen in Figures 1, 2 and 3. Very few dots are placed along straight line borders, but along irregular shaped borders points are traced very close to each other. This is due to the way in which the shapefiles are authored. As a result the geonear command may not always calculate the shortest distance between two points. As can be seen from figure 2 and 3 this occurs on over a short distance on the Angolan and Zambian borders, but these effects are considered to be negligible overall.

**Figure 2: Points along Namibia-Angola border**

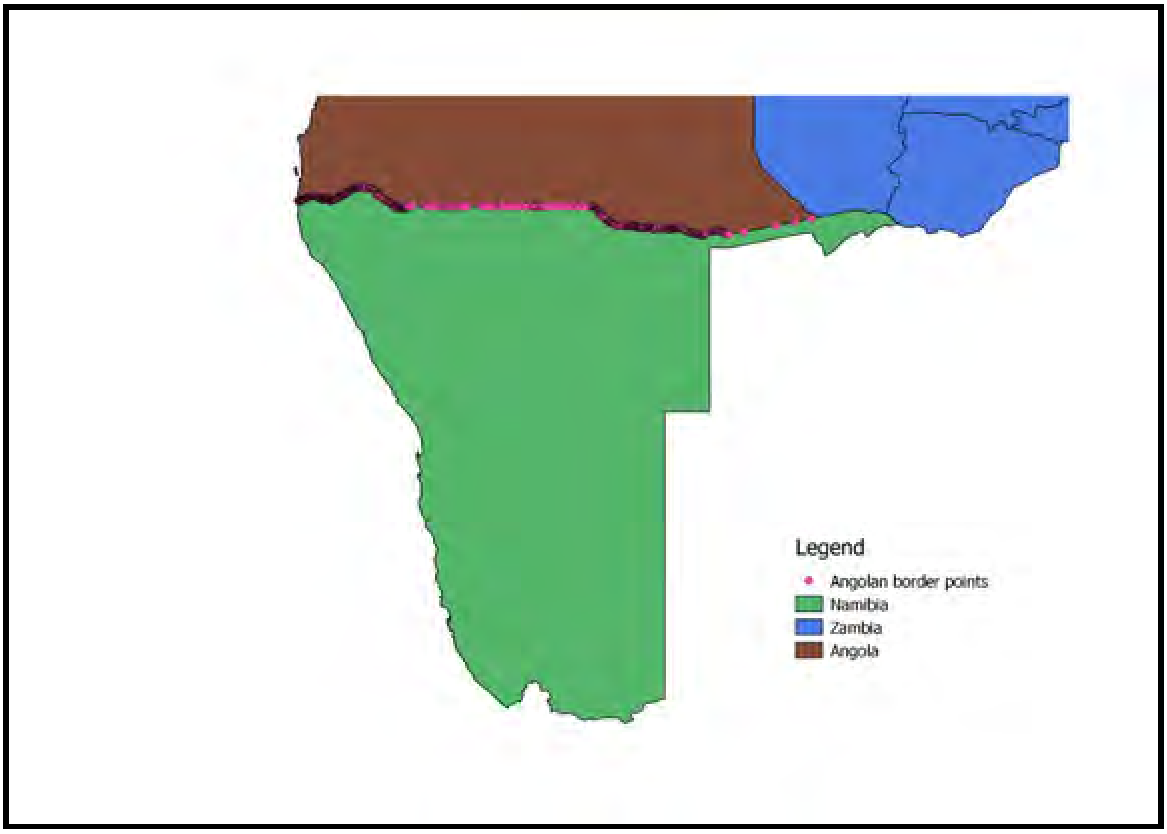


**Figure 3: Points along Namibia-Zambian border**

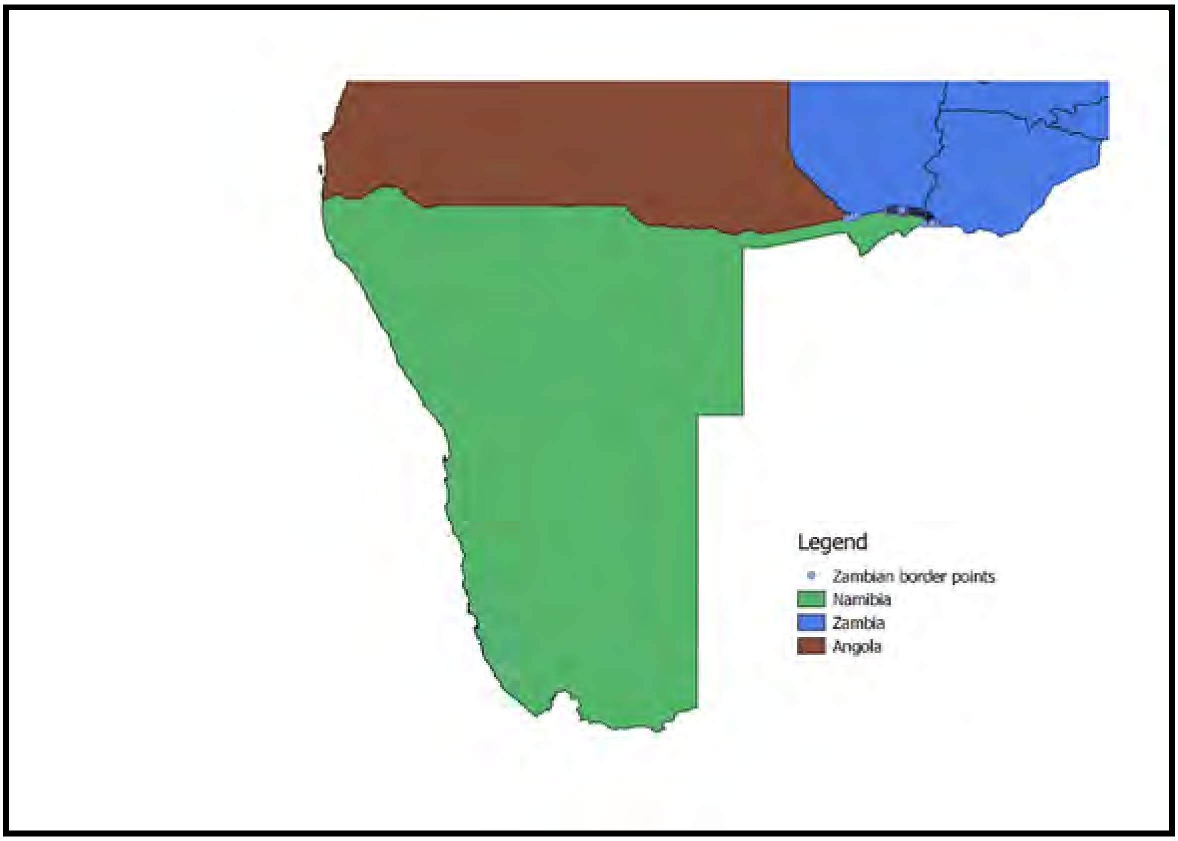
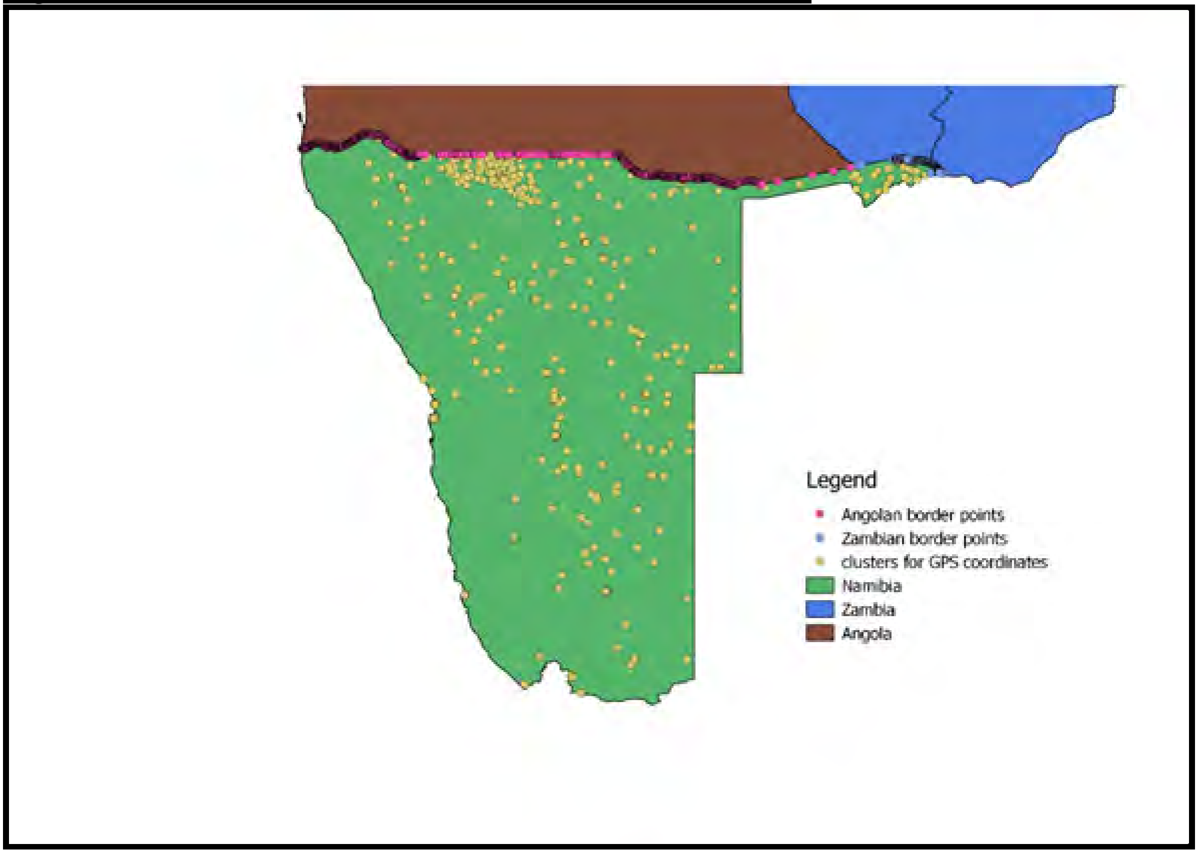
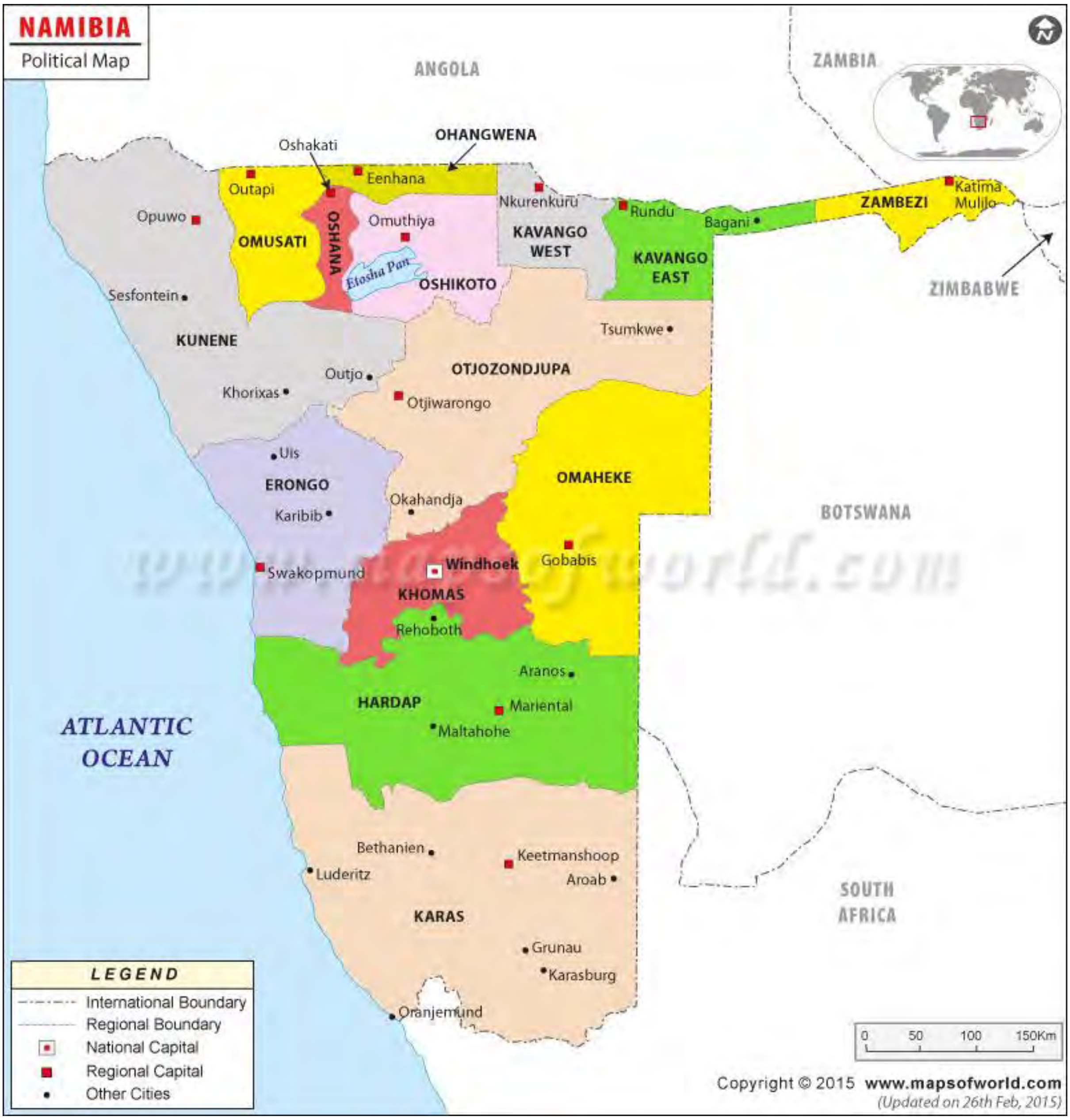



**Figure 4: GPS coordinates of clusters in relation to borders**



# Appendix E: Political map of Namibia




Source: Maps of the World, 2015