Healthy communities - Understanding determinants of HIV

12 December 2014

1 Introduction

The expiry date of the Millennium Development Goals (MDGs) is just around the corner, meanwhile the post-2015 agenda is being discussed intensively. In this context, it is important to assess the achievement of the MDGs and try to understand why some goals have not been reached.

Reducing HIV prevalence is an important aim of the MDGs. Target 6.A of the MDGs specifies that countries should "have halted by 2015 and begun to reverse the spread of HIV/AIDS" (United Nations 2014). In most regions of the world this goal has been fulfilled: new HIV infections declined and the overall number of new HIV/AIDS infections per 100 adults (15-49 years old) decreased by 44 per cent between 2001 and 2012 (United Nations 2014). However, this trend cannot be observed in all 189 member states of the United Nations. On the contrary, HIV/AIDS prevalence has even increased in some countries.

---

2 Aim, Research Question and Hypotheses

This paper aims to provide evidence to assess why some countries struggle to achieve MDG 6A. We believe that one possible explanation for the failure of some interventions in reducing HIV/AIDS may lie in the lack of a full understanding of the determinants of the disease, which can in turn lead to ill-specified interventions and wrongly targeted campaigns.

The literature reviewed for this paper identifies a myriad of determinants of health. However, there is a gap in the literature when it comes to the determinants of specific diseases. The aim of this paper is to test the applicability of one of the most commonly used theories of determinants of health, to explain the evolution of HIV/AIDS rates. The first hypothesis of this paper is that all general determinants of health (as identified by our reference model) are determinants of HIV/AIDS. By identifying variables that help to explain HIV/AIDS incidence and prevalence, this paper will help move forward the discussion of the determinants of HIV/AIDS .

Furthermore, given that the literature identifies interlinkages between the different determinants of health (see Solar and Irwin (2010)), this paper will evaluate to what extent those linkages are reflected on the micro level. The second hypothesis of this paper is that socioeconomic, environmental and cultural factors can be used to explain individual lifestyle factors.

---

3 Literature Review

Hurrelmann (Hurrelmann 1989, 76) advocates an interdisciplinary framework for analysing what determines health outcomes. He considers it necessary to use a model that integrates all the aspects of the organism, individual and the environment.

One framework that shows the interaction between individual and environmental factors over time is the salutogenic model developed by Antonovsky in 1979. According to Hurrelmann, Antonovsky's model is a great contribution to interdisciplinary theory, but the downside is its complexity (Hurrelmann (1989)).

A simpler and more common model on the main determinants of health is the “rainbow model”, developed by Dahlgren and Whitehead (Dahlgren and Whitehead (1991), p.11). This model gives an overview of the main health determinants, reflecting the relationship between the individual, its environment and different health outcomes. Individuals are at the centre of the model with a set of fixed biological and genetical preconditions. Building upon these, four layers of influence on health can be identified: individual lifestyle factors, social and community networks, living and working conditions and general socio-economic, cultural and environmental conditions.

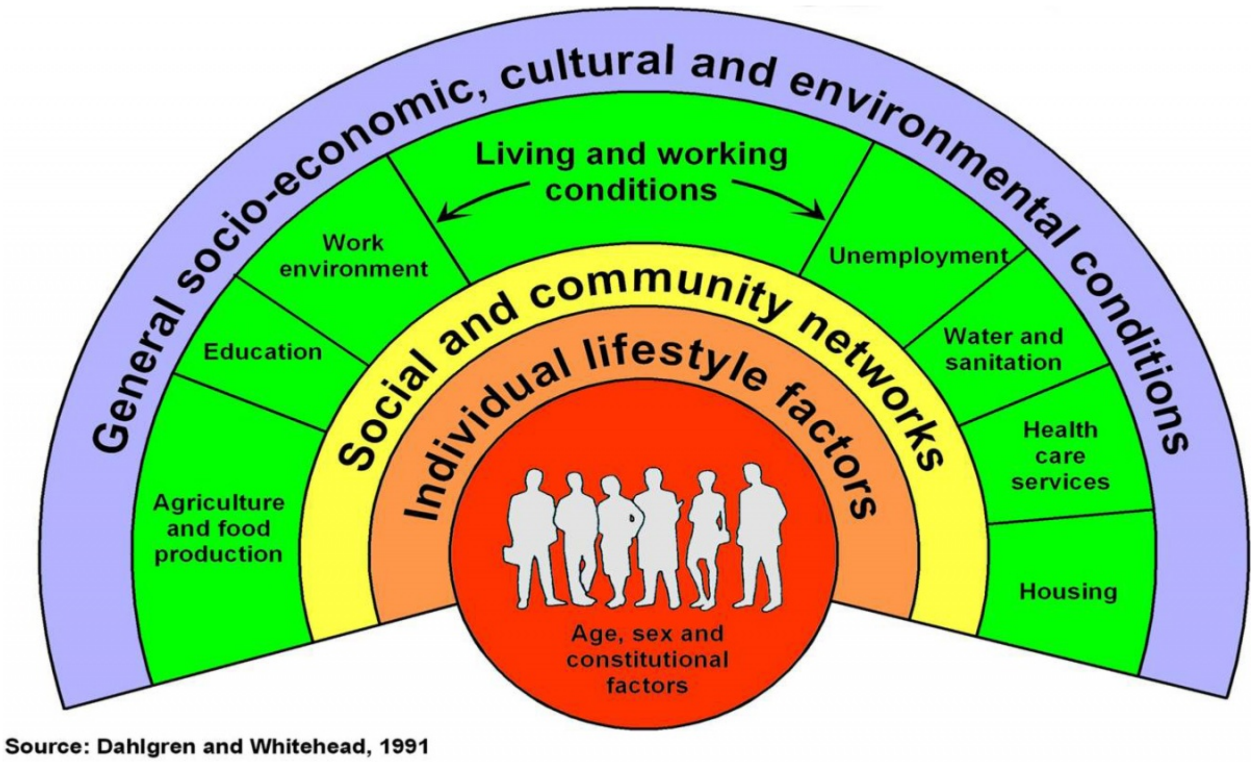


Figure 1: Dahlgren and Whitehead (Dahlgren and Whitehead (1991), p.11). Main health determinants.

---

4 Data Sources

For this research, two main sources of data were explored. Firstly, databases comprising information for several countries were examined in order to increase comparability of the data. Among the most commonly used databases in the field of this research, it is worth mentioning those from UNAIDS, World Bank, Global Fund for AIDS, Tuberculosis and Malaria, WHO, the Institute for Health Metrics and Evaluation, PEPFA and the AIDS Data Hub. Secondly, country-specific data was used to check the trends of some of the most important variables of this research and to enable micro level analysis. These databases include those from the National Bureau of Statistics (from each country under analysis), Demographic Health Surveys (DHS) and the AIDS Indicator Survey (AIS), obtained from the United States Agency for International Development (USAID).

4.1 World Development Indicators (WDI)

The WDI database comprises 1342 indicators clustered in 10 thematic areas that range from health and education to infrastructure and public sector data. Information is available for 214 countries and dates back to 1960. All indicators are available for free at the World Bank website and can be downloaded as an Excel sheet, CSV, tabbed TXT or SDMX. In addition, there is a special R package WDI designed to download and use the data.

WDI have been used in a wide range of fields and HIV/AIDS research is not an exception. For examples of relevant literature that also make use of WDI please see Haacker (2002), Talbott (2007) and Kalemli-Ozcan (2011). A list of all WDI indicators used for this research can be found in the Appendix.

4.2. Demographic Health Surveys (DHS) and AIDS Indicator Survey (AIS)

DHS have a high reputation for collecting accurate and nationally representative household data (World Bank (2014)). Relevant HIV/AIDS literature used the DHS data like Haacker (2002) and World Health Organization and UNICEF (2014). For the case of Vietnam three datasets for the years 2005, 2002 and 1997 are available in USAID's DHS database. The most recent available DHS dataset will be used for the micro level analysis of this research. The DHS was conducted by the General Statistical Office National Institute for Hygiene and Epidemiology between September and December 2005. There are three core questionnaires in DHS surveys: a household, a women's and a men's questionnaire. The DHS sample for Vietnam counts 6337 households, including 7289 observations for female respondents (women aged 15 to 49) and 6707 observations for male respondents (men aged 15 to 49).

As an extension of the DHS data, the micro analysis also integrates the AIDS Indicator Survey (AIS) for Vietnam, again being conducted in 2005. The AIS provides nationally representative estimates of HIV rates, by collecting blood from representative samples of the population of both men and women in a country. The AIS for Vietnam includes 1675 households. In all households, women and men aged 15-49 are eligible to participate in the survey. The linkage of DHS' HIV test results to the full DHS survey record (without personal identifiers) allows for an in-depth analysis of the socio-demographic and behavioural factors associated with HIV prevalence.

Both datasets can be downloaded as Stata System File, Flat Data, SAS System File and SPSS System File. Individual and AIS datasets can also be downloaded as a hierarchical file format.

It is worth mentioning that special permission was granted in downloading the DHS and AIS databases, representing a potential challenge for the reproducibility of our research.

---

5 Data Gathering and Cleaning

This section focuses on the process of gathering the data and cleaning the databases to prepare the variables for the data analysis.

The first step in this process was uploading the databases to R Studio. The first dataset consists of 29 World Development Indicators and it was downloaded from World Bank's website. These indicators represent the independent variables used for this research plus the population indicator that is used to filter small countries. Provided that the focus of this research is on country level data, all regional data was dropped. Further, 169 rows that contained only NA values were deleted.

After dropping empty rows, the data frame was alphabetically (ascending) ordered, rows were grouped by iso2c code and variables were renamed.

The dataset was further cleaned preparing the data for imputation using the AMELIA package. The imputation will be conducted however at a further stage of the research. This process requires that the panel is as balanced as possible, as it feeds from all variables to predict values for the missing observations. The next step was thus dropping variables for which more than 80% of the observations (552) were missing. In addition, countries with a population smaller than one million inhabitants were dropped from the database. 59 countries fell in that category: 46 islands, 5 European countries (Andorra, Liechtenstein, Luxemburg, Monaco and Montenegro), Bahrain, Bhutan, Belize, Djibouti, Equatorial Guinea, Guyana, Qatar and Suriname. Dropping these countries does not affect the research as the remaining database still contains a highly heterogeneous sample both in geographic and socio-economic terms. Furthermore, deleting these countries improves the dataset as most of these countries lacked information for most of the studied variables.

The second database used for this research was downloaded from UNAIDS' website and it provides information on HIV/AIDS incidence rates (as well as prevalence and deaths caused by HIV/AIDS). The data is publicly available. All columns except the country and the incidence rate were dropped. After renaming the variables, a unique identifier was created and missing values were recoded as NAs. Moreover, some observations in the database were not specific numbers; instead, it was indicated that for that year, prevalence was below a certain threshold (0.01%). In those cases, these observations were replaced by 0.009. The final step in the cleaning of the UNAIDS database consisted of deleting missing values for the dependent variable and deleting the regions with an iso2c equal to a country's iso2c (NA and ZA) to avoid problems in the merging process.

Once both databases were cleaned, the next step was to merge the datasets using the combination of iso2c and year as unique identifier. In the merging process, only observations that were present in both datasets were kept. It is worth noticing that UNAIDS' dataset included observations from 1990 to 2012 so all observation between 1990 and 1999 were dropped. Finally, unnecessary columns from the new database were eliminated.

6 Descriptive Statistics

The descriptive statistics part consists of the preparation of the variables for data analysis. Tables, plots and histograms are shown to understand the distribution of the variables.

The histogram of the dependent variable (Figure 1) shows that the incidence rates are not normally distributed but strongly skewed to the left and only few incidence rates are higher than 1.

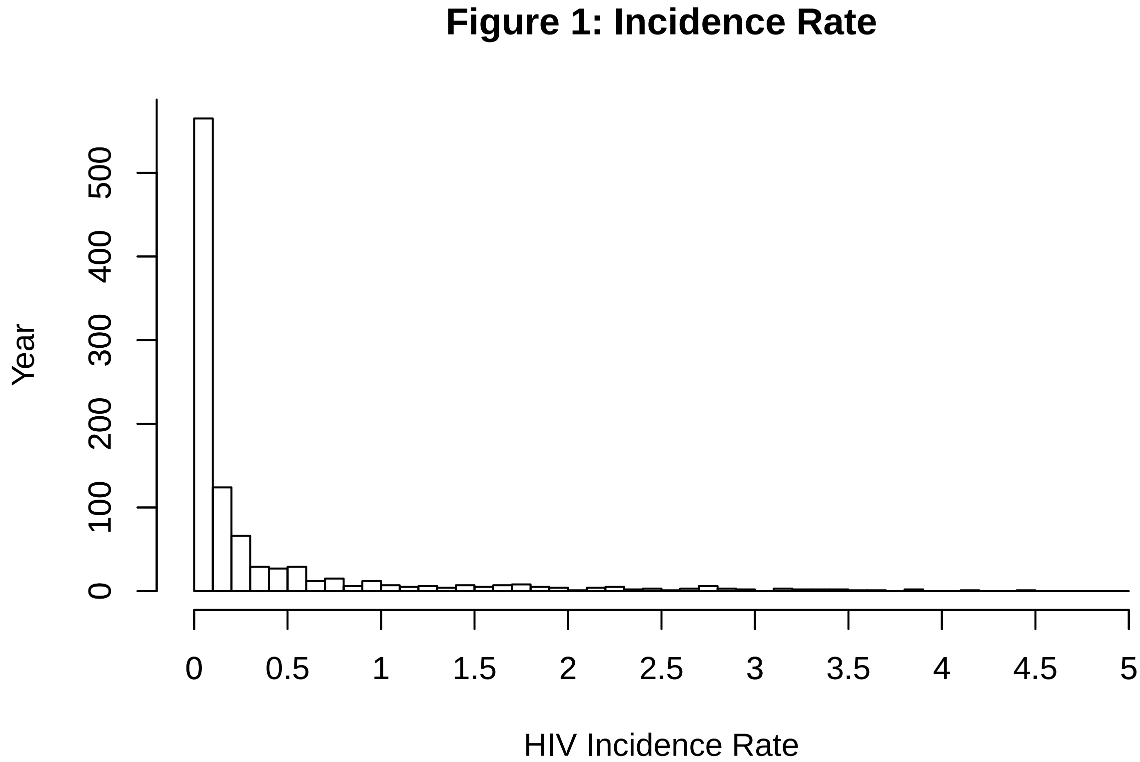
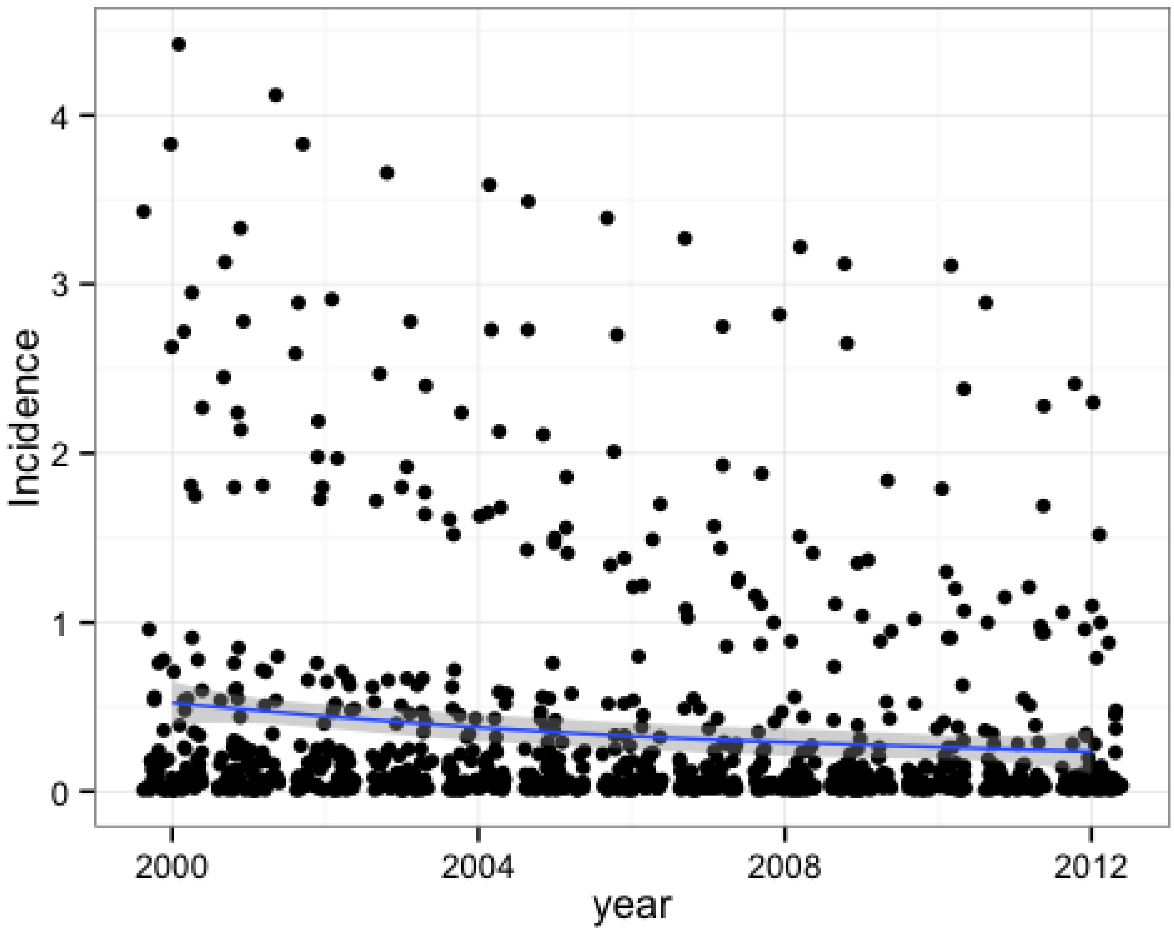


Figure 2 shows that in most countries of our dataset HIV/AIDS incidence rates decreased between the period of 2000 to 2015 (see Figure 2).

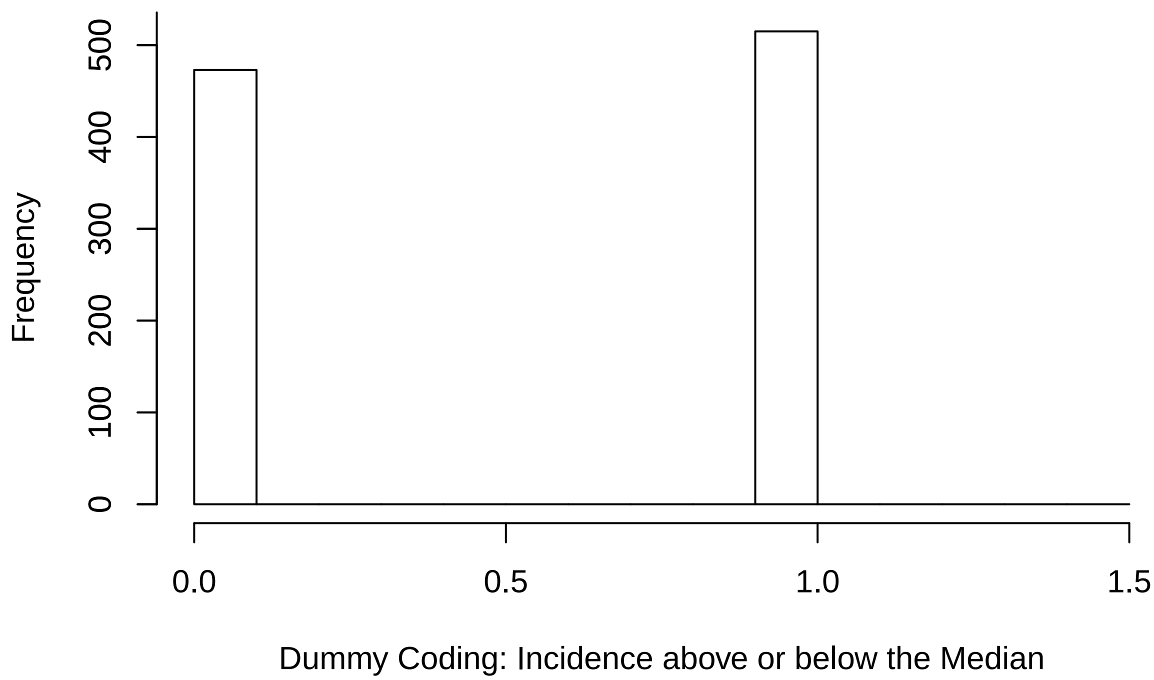


When plotting the incidence rates per country (Figure 3 in appendix) the range of observations per country is shown and the outliers (countries with high incidence rates) can be identified.

As the research question investigates why MDG 6.A is not being reached by some countries, the general HIV incidence rate is interesting, but also the decrease in the incidence rate from 2000 to 2015 is even more relevant. As stated in the research proposal Target 6.A of the MDGs specifies that countries should "have halted by 2015 and begun to reverse the spread of HIV/AIDS" [United Nations (2014)].

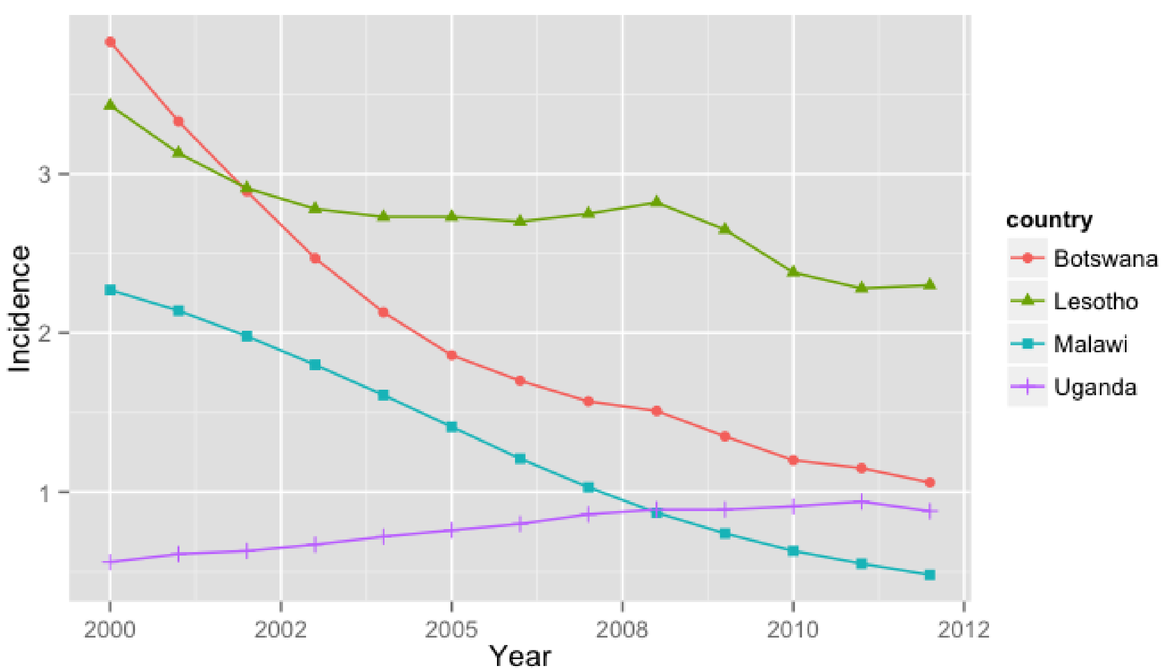
For this purpose, the dependent variable was lagged by one period and the difference between the lag and the current year was calculated (see Figure 4 in Appendix). Further, a dummy variable was created assigning a value of zero for those observations where the incidence rate decreased compared to the previous year or stayed the same (countries reaching MDG 6.A) and a value of one was assigned to those observations where the incidence rate increased (countries not reaching MDG 6.A). Figure 4 shows the direction of the change in the incidence rate compared to the previous year by country.

Figure 5: Dummy Variable

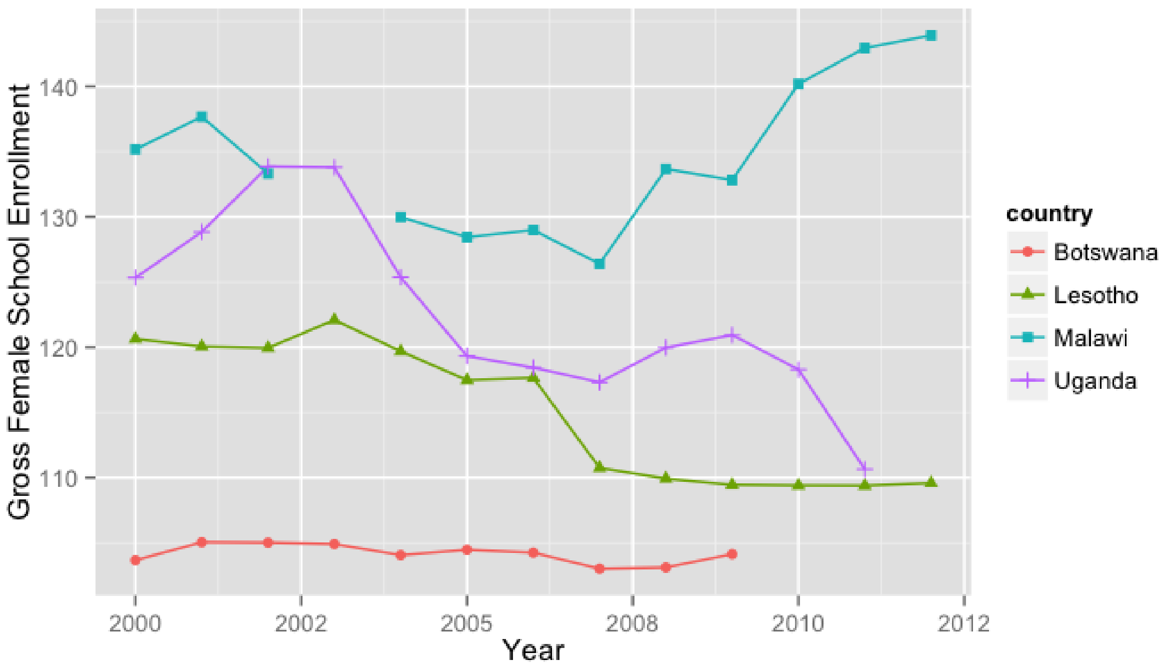


7 Case Studies - Botswana, Lesotho, Uganda & Malawi

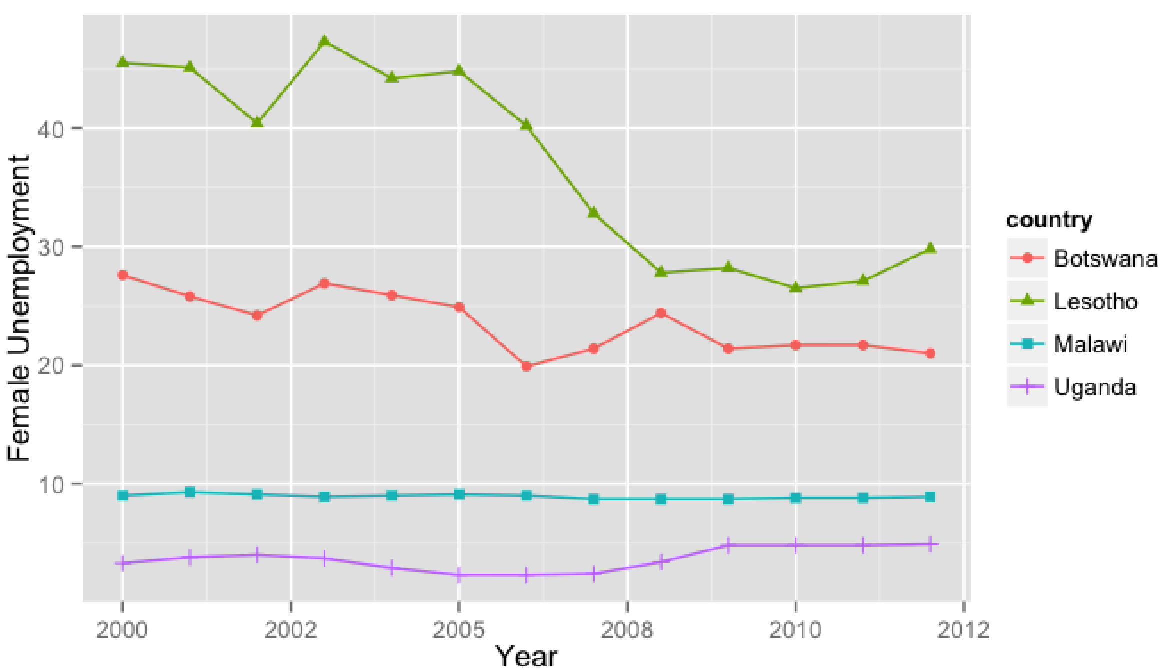
7.1 Case Studies



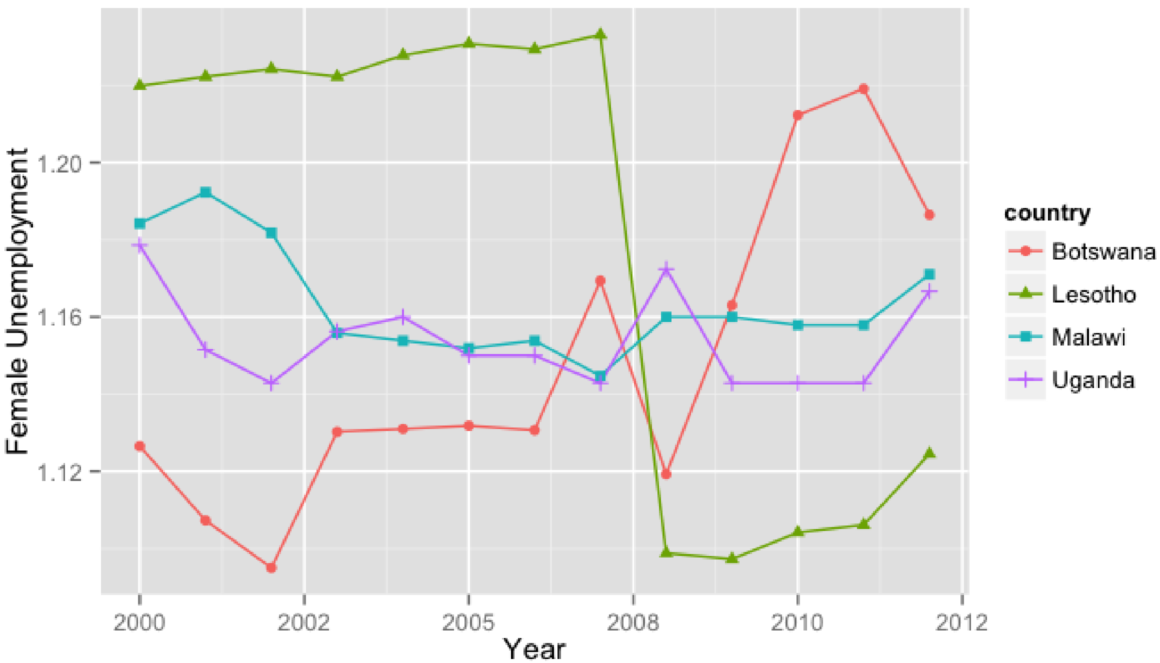
7.2 Female Schooling in Selected Countries



7.3 Female Unemployment in Selected Countries



7.4 Female Unemployment Share in Selected Countries



Scatterplots were used for each category of Dahlgrehn's model in order to see whether the variables are skewed or multicollinear (see Figures 5, 6 & 7 in Appendix).

8 Inferential Statistics

As can be seen from the Scatterplots (see appendix) most of the variables are not normally distributed. Further, the variables all have different scales. Therefore, the independent variables were logged for enabling comparisons in the data analysis part.

8.1 Model 1 - Comparison of countries with an HIV Incidence Rate below and above the Median

For Model 1 logistic regressions are used for predicting the likelihood that a country has a low HIV incidence rate (the dependent variable Y is equal to 1, rather than 0) given certain values of the HIV determinants (values of X) being hypothetically assumed.

As the incidence variable is highly skewed to the left, the most accurate measure of the central tendency of the distribution is the median. Therefore, the dependent variable is coded as a dummy variable, being 1 if the HIV incidence rate is below the median and 0 for countries with an HIV incidence rate above the median.

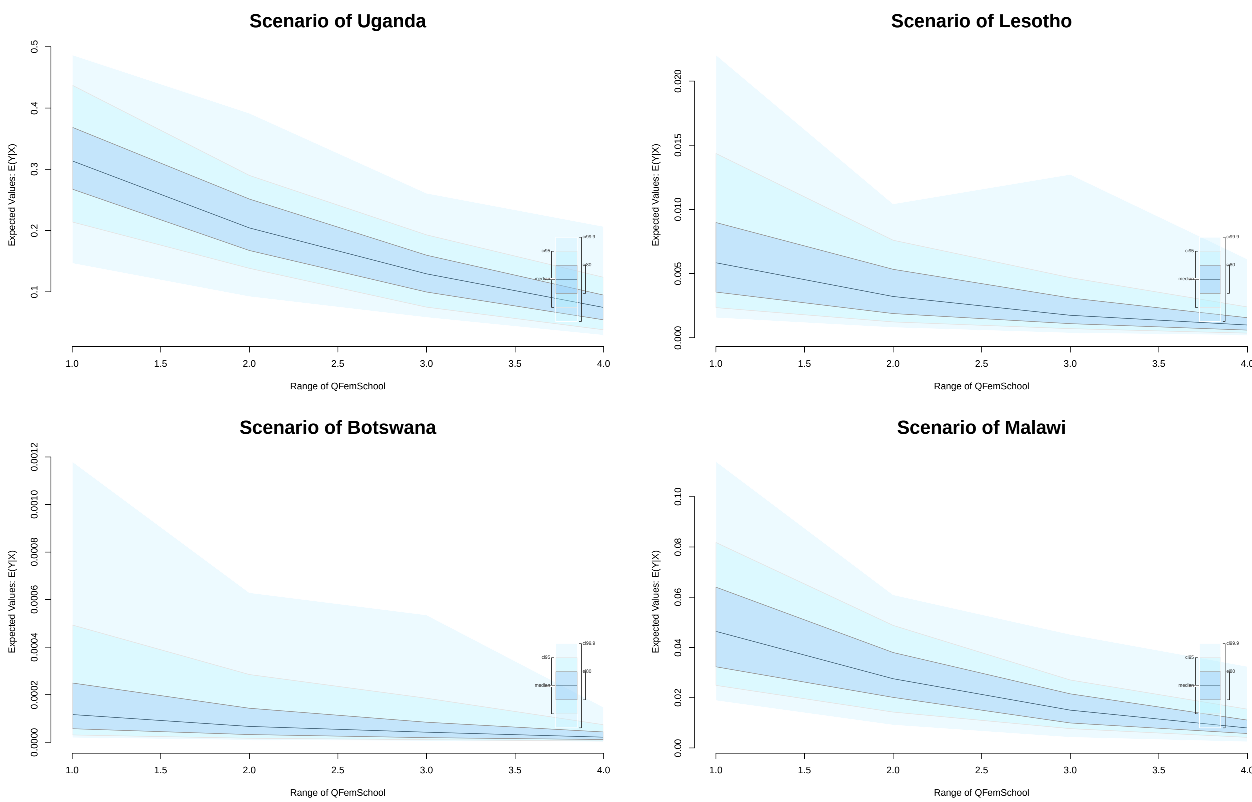
As Odds and Odds ratios are difficult to interpret and to present to a broad audience predicted probabilities are calculated after running the logistic regressions. The interpretation of the logistic regression results presented in table 1 will focus only on the significance of the variables.

A more detailed analysis of the predicted probabilities will be given for the main variables of interest Female School Enrollment and Share of Female Unemployment. Predicted probability scenarios are calculated for the four case studies that were introduced in the descriptive statistics part.

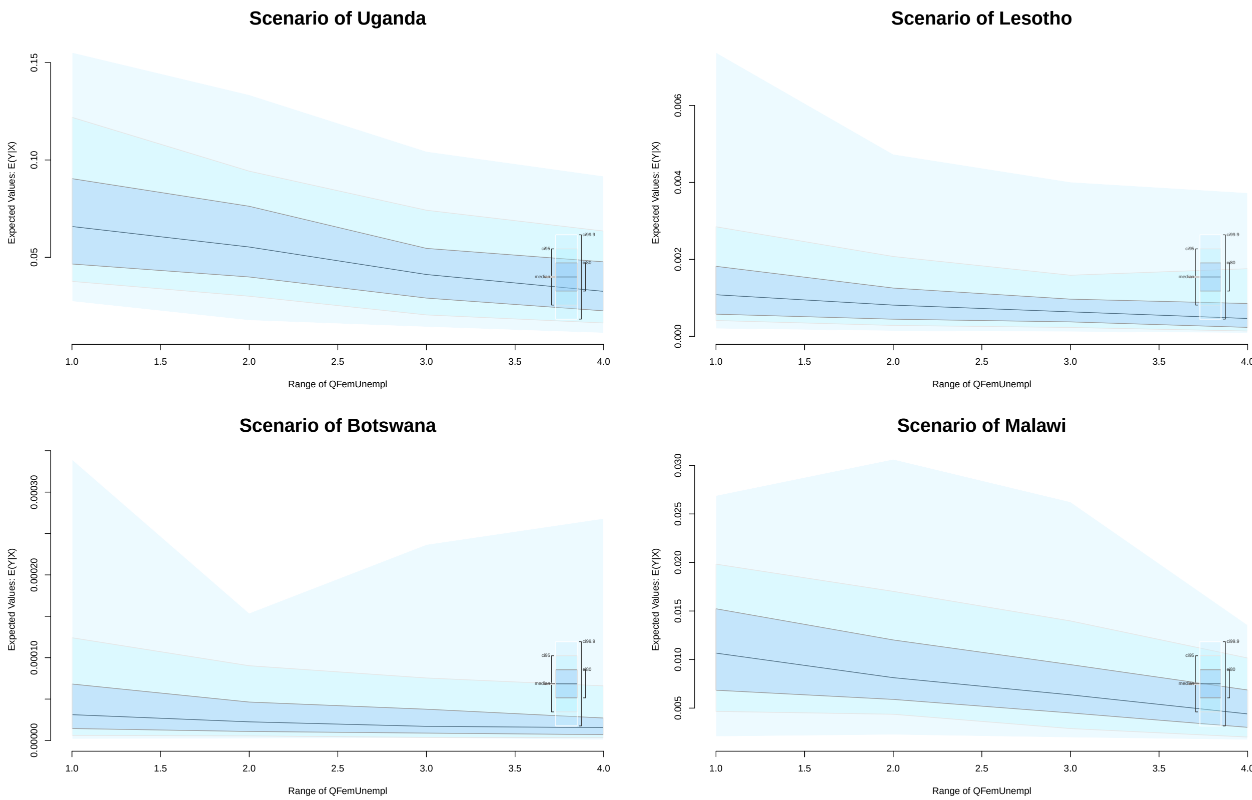
Table 1: Logistic Regression Results of Model 1

Variables	Coefficients	Std. Error	T-Statistic	P-Value
Constant	-102.47	9.74	-10.52	0.00
GDP per capita	-0.78	0.33	-2.39	0.02
Share of Rural Population	-1.03	0.41	-2.48	0.01
Share of CO2	-1.01	0.19	-5.37	0.00
Healthcare Expenditure	0.34	0.38	0.89	0.38
Access to Water	0.30	0.76	0.39	0.70
Access to Sanitation	0.24	0.24	1.02	0.31
Life Expectancy	30.26	2.53	11.97	0.00
Immunisation against DPT	-1.18	1.15	-1.03	0.30
Immunisation against Measles	1.42	1.22	1.17	0.24
Female School Enrollment	-3.54	0.54	-6.50	0.00
Share of Female Unemployment	-0.43	0.50	-0.86	0.39

Predicted Probabilities for Female School Enrollment



Predicted Probabilities for Share of Female Unemployment



8.2 Model 2 - Focusing on Countries with an HIV Incidence Rate above the Median

Table 2: OLS Regression Results of Model 2

Variables	Coefficients	Std. Error	T-Statistic	P-Value
Constant	16.27	1.57	10.34	0.00
GDP per capita	0.21	0.08	2.70	0.01
Share of Rural Population	0.48	0.12	3.84	0.00
Share of CO2	0.09	0.04	2.10	0.04
Healthcare Expenditure	0.06	0.10	0.57	0.57
Access to Water	0.51	0.19	2.64	0.01
Access to Sanitation	0.01	0.07	0.16	0.87
Life Expectancy	-7.50	0.30	-24.73	0.00
Immunisation against DPT	0.30	0.27	1.12	0.26
Immunisation against Measles	-0.17	0.29	-0.57	0.57
Female School Enrollment	1.36	0.16	8.37	0.00
Share of Female Unemployment	1.46	0.19	7.78	0.00

Table 3: OLS Regression Results of Model 2 with robust standard errors

Variables	Coefficients	Std. Error	T-Statistic	P-Value
Constant	16.27	1.57	10.34	0.00
GDP per capita	0.21	0.08	2.70	0.01

Variables	Coefficients	Std. Error	T-Statistic	P-Value
Share of Rural Population	0.48	0.12	3.84	0.00
Share of CO2	0.09	0.04	2.10	0.04
Healthcare Expenditure	0.06	0.10	0.57	0.57
Access to Water	0.51	0.19	2.64	0.01
Access to Sanitation	0.01	0.07	0.16	0.87
Life Expectancy	-7.50	0.30	-24.73	0.00
Immunisation against DPT	0.30	0.27	1.12	0.26
Immunisation against Measles	-0.17	0.29	-0.57	0.57
Female School Enrollment	1.36	0.16	8.37	0.00
Share of Female Unemployment	1.46	0.19	7.78	0.00

Table 4: Fixed Effects Regression Results of Model 2

	Value	Std. Error	t-stat	p-value
(Intercept)	-2.6899	2.4815	-1.0840	0.2785
lGDPpc	0.0124	0.1077	0.1147	0.9087
lRural	3.1506	0.3592	8.7721	0.0000
lCO2	0.0205	0.0309	0.6653	0.5061
lHCexpend	-0.1523	0.0767	-1.9852	0.0486
lWater	-1.1798	0.3296	-3.5796	0.0004
lSanitation	-0.4045	0.2502	-1.6169	0.1069
lLifeExpect	-1.1979	0.3706	-3.2320	0.0013
lDPT	0.2544	0.1214	2.0950	0.0362
lMeasles	-0.3950	0.1269	-3.1119	0.0019
lFemSchool	-0.0317	0.1256	-0.2528	0.8035
lShFemUnempl	-0.0557	0.1699	-0.3281	0.7428
as.factor(country)Benin	-0.7106	0.4430	-1.6043	0.1097
as.factor(country)Botswana	3.8813	0.2444	15.8832	0.0000
as.factor(country)Burkina Faso	-1.8621	0.3783	-4.9224	0.0000
as.factor(country)Burundi	-0.7585	0.3227	-2.3506	0.0187
as.factor(country)Cambodia	-2.1160	0.3346	-6.3238	0.0000
as.factor(country)Cameroon	1.8231	0.1885	9.6717	0.0000
as.factor(country)Central African Republic	0.3857	0.3346	1.1526	0.2496
as.factor(country)Chad	-1.2071	0.3724	-3.2411	0.0013
as.factor(country)Congo, Dem. Rep.	-0.9202	0.2623	-3.5082	0.0005
as.factor(country)Congo, Rep.	1.5056	0.4743	3.1743	0.0016
as.factor(country)Cote d'Ivoire	0.6874	0.3486	1.9718	0.0492
as.factor(country)Dominican Republic	2.1007	0.2309	9.0997	0.0000
as.factor(country)El Salvador	1.2646	0.3081	4.1045	0.0000
as.factor(country)Eritrea	-1.9432	0.3882	-5.0058	0.0000
as.factor(country)Ethiopia	-2.3908	0.3873	-6.1726	0.0000
as.factor(country)Gabon	5.4496	0.5460	9.9817	0.0000
as.factor(country)Gambia, The	1.3847	0.2486	5.5692	0.0000
as.factor(country)Ghana	0.2507	0.4556	0.5503	0.5825
as.factor(country)Guatemala	0.6477	0.1810	3.5785	0.0003
as.factor(country)Guinea-Bissau	0.7923	0.3633	2.1808	0.0301
as.factor(country)Haiti	0.4042	0.3014	1.3414	0.1810
as.factor(country)Honduras	0.4823	0.2419	1.9937	0.0462
as.factor(country)Jamaica	1.8105	0.1899	9.5346	0.0000
as.factor(country)Kenya	0.2870	0.2279	1.2593	0.2084
as.factor(country)Lesotho	2.1345	0.3201	6.6680	0.0000

	Value	Std. Error	t-stat	p-value
as.factor(country)Liberia	0.0405	0.4367	0.0928	0.9262
as.factor(country)Madagascar	-2.1956	0.4063	-5.4033	0.0000
as.factor(country)Malawi	0.5711	0.5085	1.1230	0.2627
as.factor(country)Mali	-1.2793	0.2829	-4.5229	0.0000
as.factor(country)Mauritius	1.2199	0.2205	5.5328	0.0000
as.factor(country)Mozambique	0.9229	0.3130	2.9484	0.0035
as.factor(country)Myanmar	-0.7585	0.2932	-2.5869	0.0098
as.factor(country)Namibia	2.3927	0.2779	8.6086	0.0000
as.factor(country)Nepal	-1.7204	0.3648	-4.7160	0.0000
as.factor(country)Niger	-2.6309	0.4517	-5.8239	0.0000
as.factor(country)Nigeria	0.5633	0.2080	2.7076	0.0072
as.factor(country)Papua New Guinea	-2.2275	0.2804	-7.9453	0.0000
as.factor(country)Rwanda	-0.2624	0.2817	-0.9314	0.3517
as.factor(country)Senegal	-0.4628	0.2110	-2.1934	0.0283
as.factor(country)Sierra Leone	-0.7400	0.4070	-1.8179	0.0706
as.factor(country)Somalia	-2.5173	0.2890	-8.7093	0.0000
as.factor(country)South Africa	4.5277	0.2317	19.5379	0.0000
as.factor(country)Swaziland	2.2821	0.1584	14.4105	0.0000
as.factor(country)Tanzania	-0.1191	0.4238	-0.2811	0.7788
as.factor(country)Thailand	-0.1049	0.2796	-0.3753	0.7075
as.factor(country)Togo	-0.3989	0.4202	-0.9493	0.3437
as.factor(country)Trinidad and Tobago	-0.4133	0.3105	-1.3311	0.1833
as.factor(country)Uganda	0.5200	0.2828	1.8385	0.0665
as.factor(country)Ukraine	2.6792	0.2600	10.3039	0.0000
as.factor(country)Vietnam	-0.7036	0.2865	-2.4563	0.0141
as.factor(country)Zambia	1.8530	0.1628	11.3807	0.0000
as.factor(country)Zimbabwe	2.2102	0.2613	8.4599	0.0000

The test for variance inflation factors showed that in our first logistic regression model six variables showed high multicollinearity and had a higher variance inflation than the threshold of 10. We tested the multicollinearity between the variables and found that there was high multicollinearity between the GDP and GDP per capital, Unemployment and Female unemployment, Primary education and female schooling. Therefore, we excluded one of these multicollinear variables for each group based on their explanatory strength for our research question, namely unemployment, primary education and GDP.

9 Limitations

The paper had to make some compromises regarding its original aim as outlined in the first research proposal. Due to the significant amount of missing values and the presence of multicollinearity, a considerable number of variables had to be dropped and could ultimately not be integrated in the logistic regression models.

The selection of these variables was not arbitrary but followed instead the theoretical framework guiding this research, i.e. Dahlgren's model. Two levels of Dahlgren's model (Social and Community Networks and Individual Lifestyle Factors) ended up underrepresented after dropping these variables. To deal with this limitation, the research will only use the theoretical framework as an instrument to guide the selection of variables but will not utilise the findings to test the validity of the model.

In terms of the data used to run the regressions, the relative high number of countries that have already halted or reversed the spread of HIV/AIDS in our sample can lead to biased results. In the next stage of the research, the effect of excluding those countries that only halted the spread will be explored.

Another shortcoming faced at this stage was the integration of figures from the descriptive statistics into the final report. A transitory solution was to save those pictures in a subfolder of the repository.

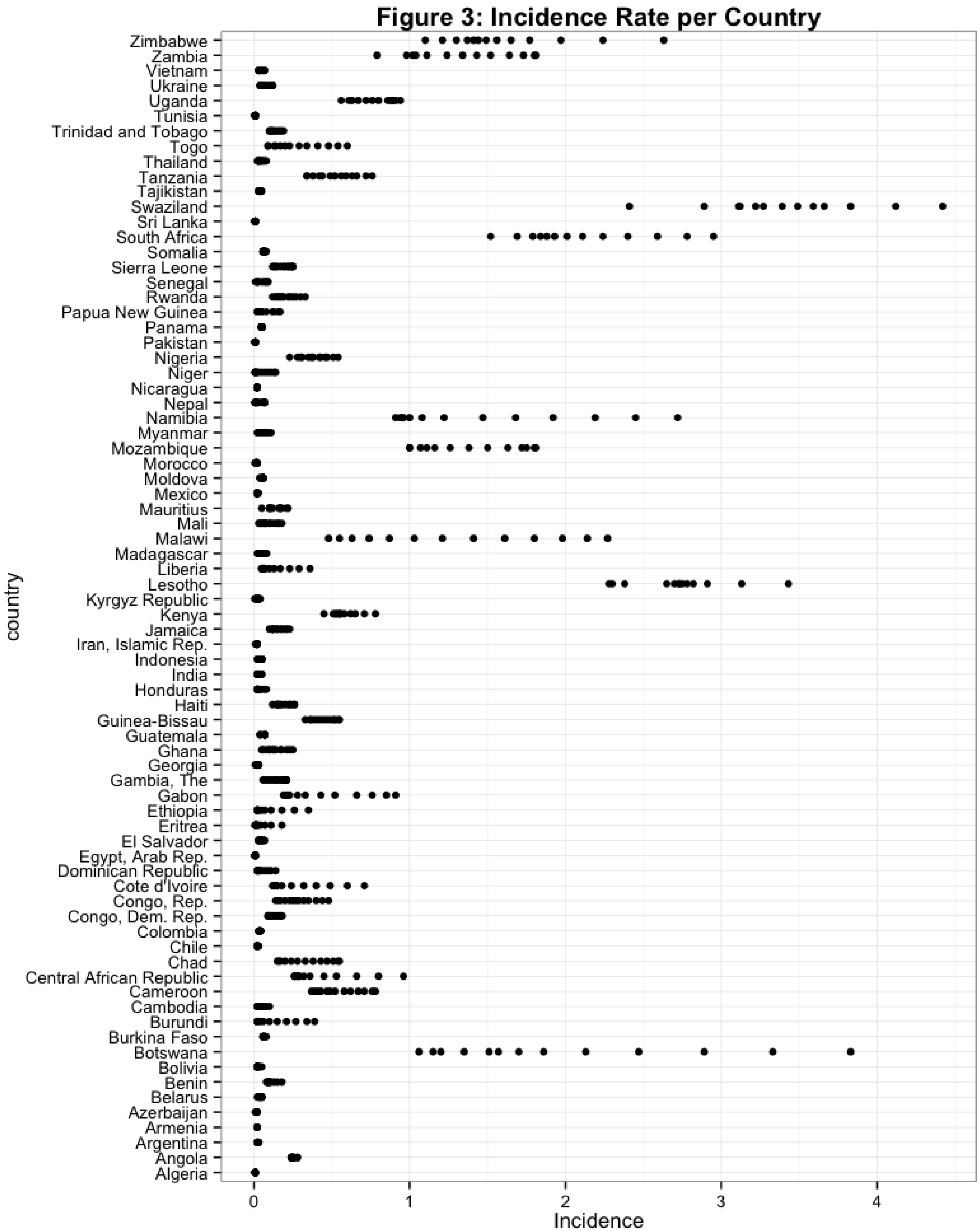
10 Appendix

Descriptive Statistics

Table 5: Descriptive statistics

Statistic	N	Mean	St. Dev.	Min	Max
X	988	494.5	285.4	1	988
year	988	2,006.0	3.7	2,000	2,012
GDPpc	945	5,296.0	5,153.0	441.2	30,875.0
Rural	988	57.7	19.5	8.7	91.8
CO2	829	1.7	3.5	0.004	38.2
HCexpend	962	5.9	2.2	1.8	18.4
Water	982	75.5	17.9	23.5	99.8
Sanitation	983	50.4	30.1	6.6	98.9
Unemploym	988	8.7	6.2	0.6	38.7
HCexpendpc	962	124.3	165.4	2.4	1,103.0
FemSchool	828	98.1	20.8	20.8	162.4
LifeExpect	988	61.4	9.9	38.1	79.6
DPT	988	79.9	18.0	19	99
Measles	988	79.4	17.5	16	99
Population	988	40,890,589.0	133,140,255.0	1,063,715	1,236,686,732
Incidence	988	0.3	0.7	0.01	4.4
IncidenceDif	912	−0.02	0.1	−0.5	0.1
DDif	912	0.9	0.3	0	1
lGDPpc	945	8.1	1.0	6.1	10.3
lRural	988	4.0	0.4	2.2	4.5
lCO2	829	−0.6	1.6	−5.6	3.6
lHCexpend	962	1.7	0.4	0.6	2.9
lWater	982	4.3	0.3	3.2	4.6
lSanitation	983	3.7	0.8	1.9	4.6
lUnemploym	988	1.9	0.7	−0.5	3.7
lHCexpendpc	962	4.1	1.2	0.9	7.0
lFemSchool	828	4.6	0.3	3.0	5.1
lLifeExpect	988	4.1	0.2	3.6	4.4
lDPT	988	4.3	0.3	2.9	4.6
lMeasles	988	4.3	0.3	2.8	4.6
lIncidence	988	−2.4	1.6	−4.7	1.5
QFemSchool	828	2.5	1.1	1	4
ShFemUnempl	988	1.1	0.3	0.5	2.5
lShFemUnempl	988	0.1	0.2	−0.8	0.9
QFemUnempl	988	2.5	1.1	1	4
Dummy	988	0.5	0.5	0	1

HIV Incidence Rates per Country over Time



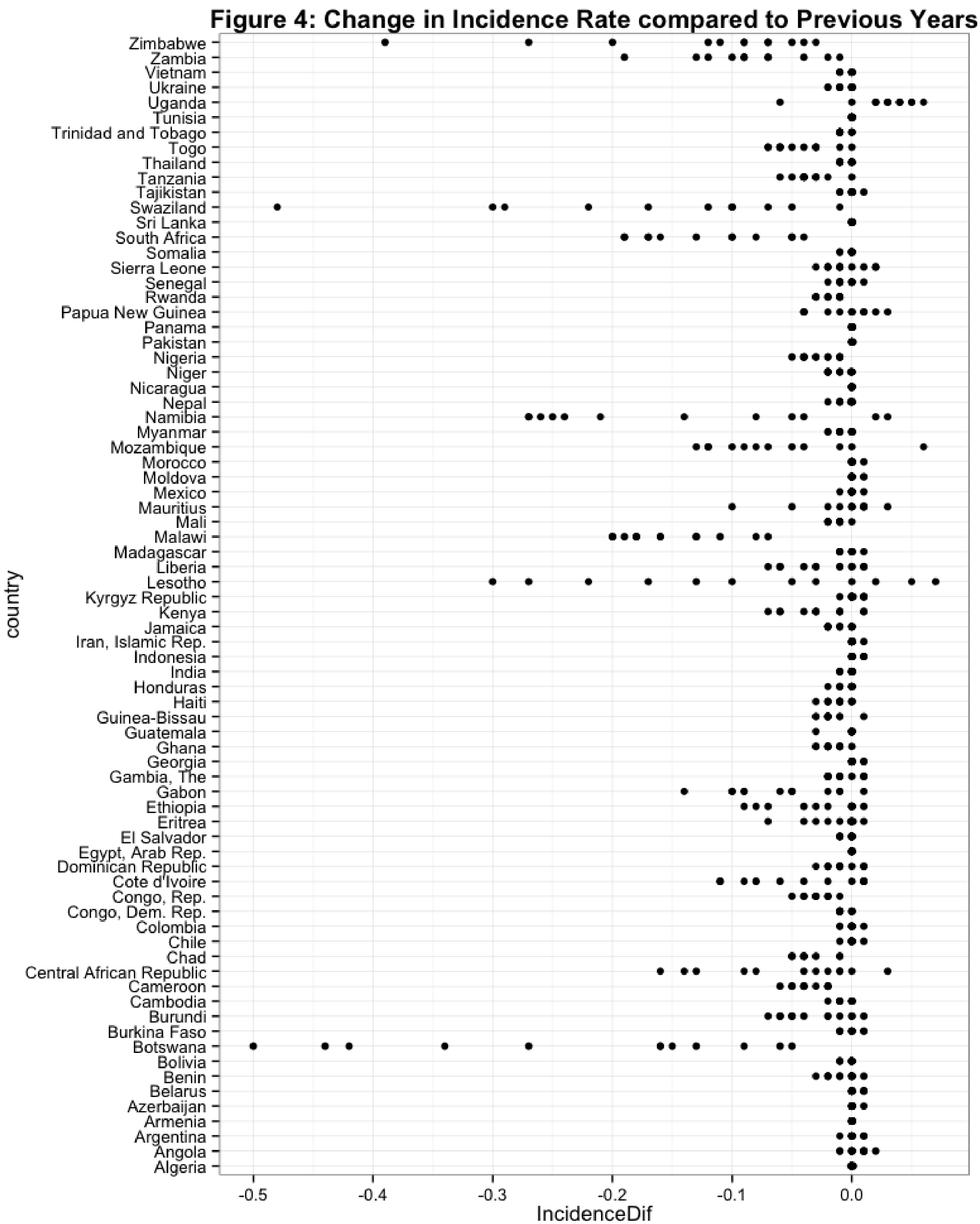
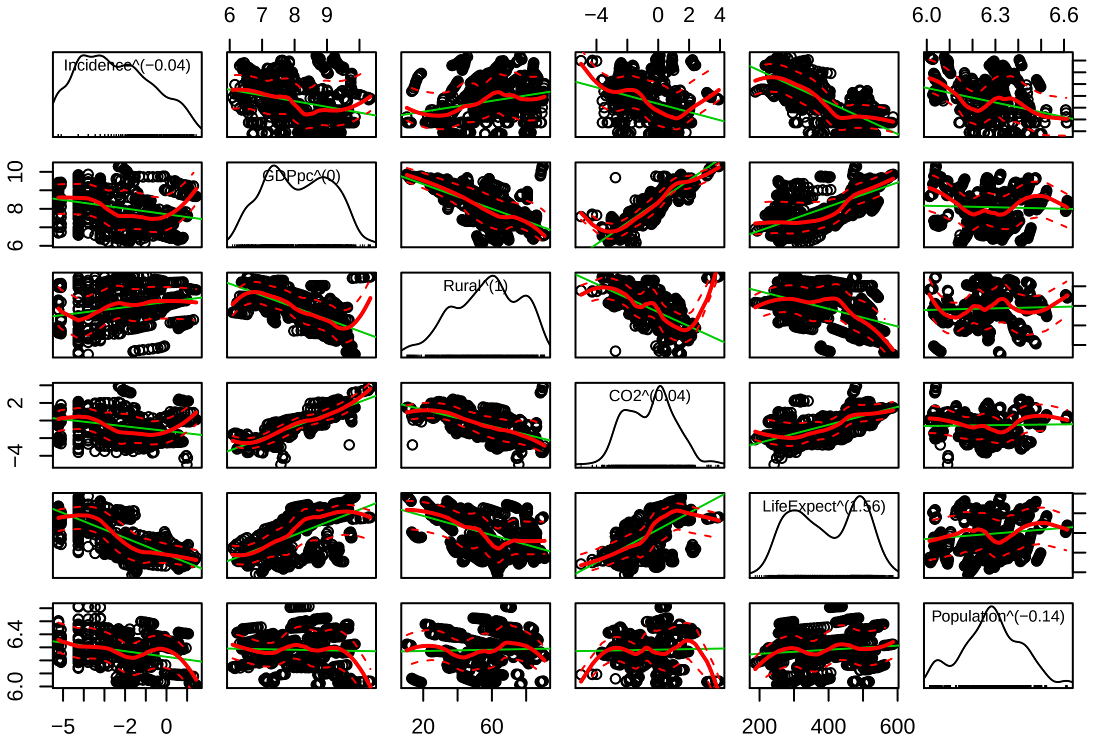
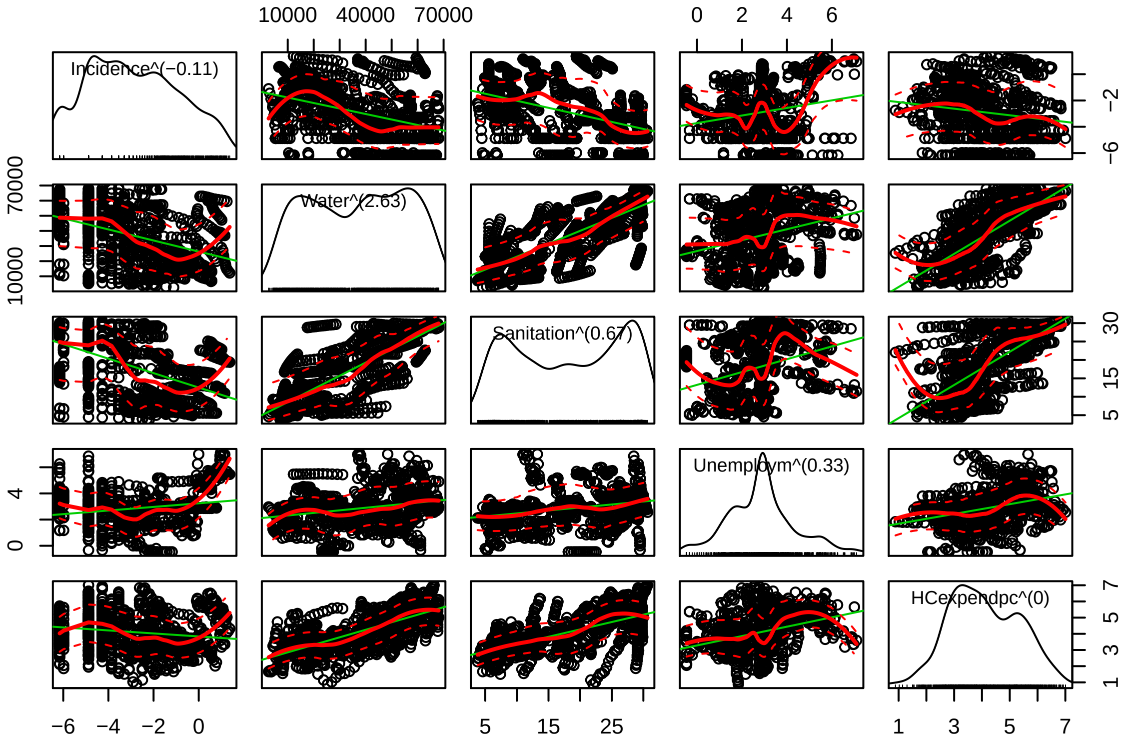


Figure 4: Change in Incidence Rate compared to Previous Years
Zimbabwe
Zambia
Vietnam
Ukraine
Uganda
Tunisia
Trinidad and Tobago
Togo
Thailand
Tanzania
Tajikistan
Swaziland
Sri Lanka
South Africa
Somalia
Sierra Leone
Senegal
Rwanda
Papua New Guinea
Panama
Pakistan
Nigeria
Niger
Nicaragua
Nepal
Namibia
Myanmar
Mozambique
Morocco
Moldova
Mexico
Mauritius
Mali
Malawi
Madagascar
Liberia
Lesotho
Kyrgyz Republic
Kenya
Jamaica
Iran, Islamic Rep.
Indonesia
India
Honduras
Haiti
Guinea-Bissau
Guatemala
Ghana
Georgia
Gambia, The
Gabon
Ethiopia
Eritrea
El Salvador
Egypt, Arab Rep.
Dominican Republic
Cote d'Ivoire
Congo, Rep.
Congo, Dem. Rep.
Colombia
Chile
Chad
Central African Republic
Cameroon
Cambodia
Burundi
Burkina Faso
Botswana
Bolivia
Benin
Belarus
Azerbaijan
Armenia
Argentina
Angola
Algeria
country
-0.5
-0.4
-0.3
-0.2
-0.1
0.0
IncidenceDif

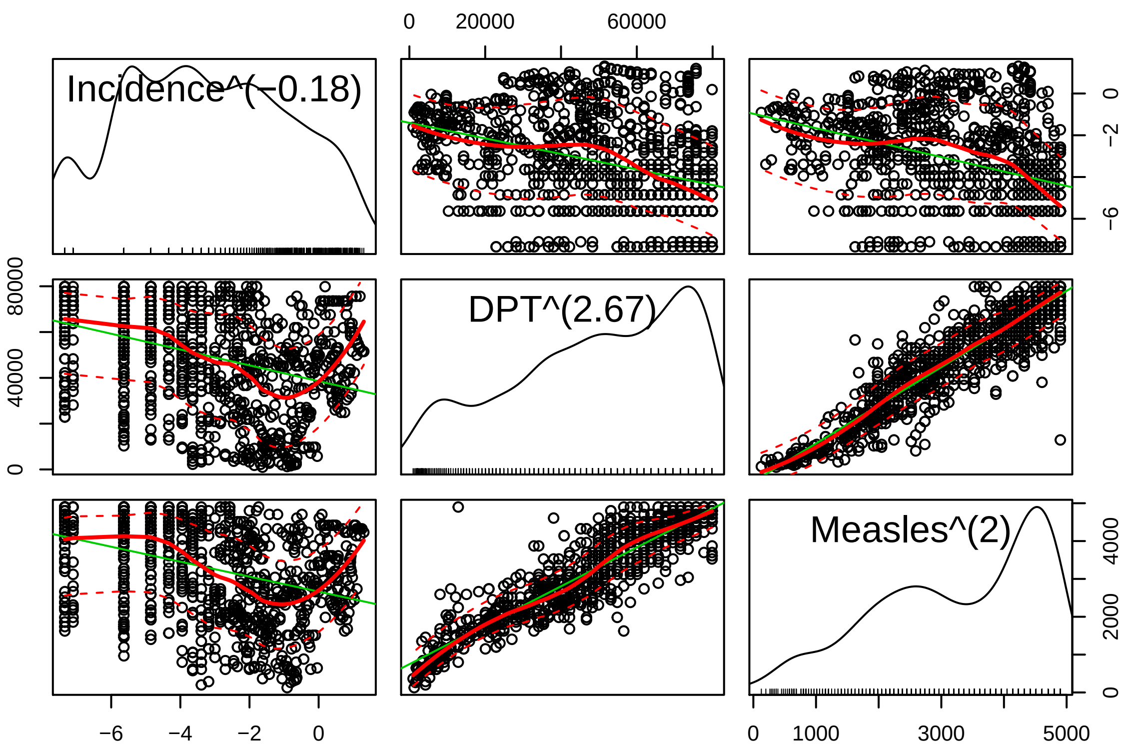
Scatterplot of variables for socio-economic, cultural and environmental conditions



Scatterplot of variables for living and working conditions



Scatterplot of variables for individual lifestyle factors



Testing for Multicollinearity

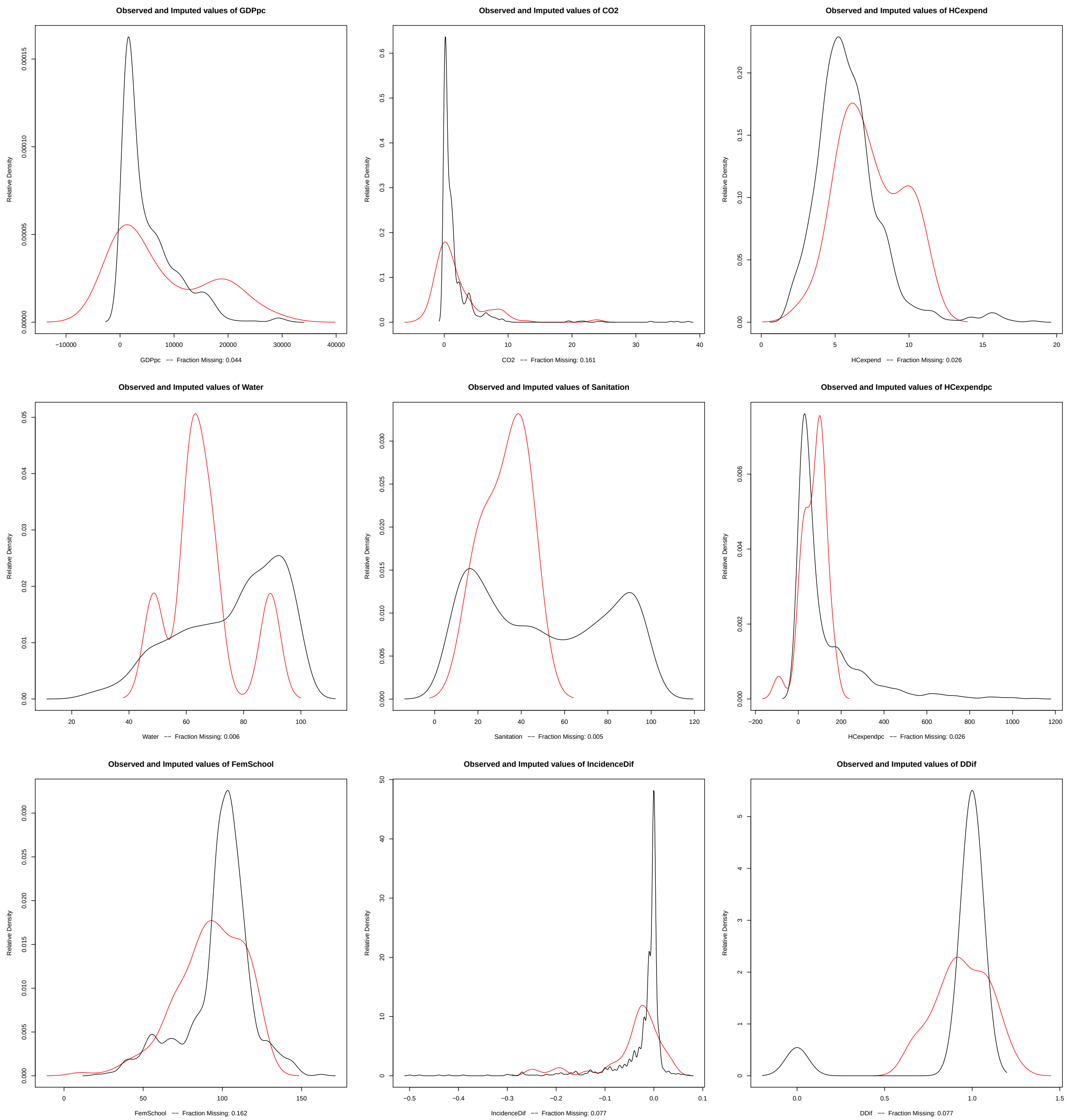
Table 6: Variance Inflation Factors - Table 1

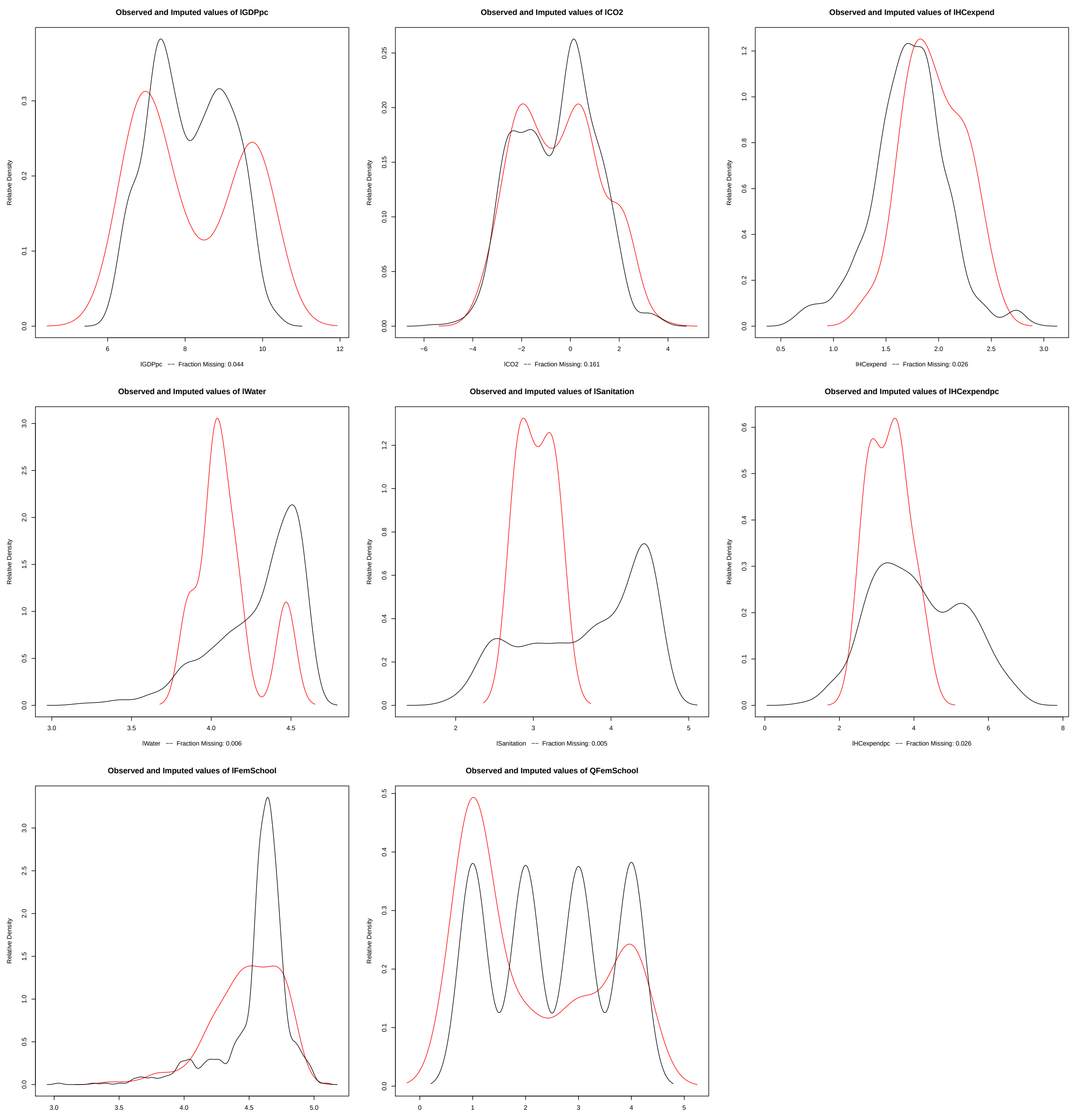
	vif
GDP	3.09
GDPpc	11.78
Rural	2.49
CO2	3.20
HCexpend	1.61
Primary	47.65
Water	3.88
Sanitation	4.41
Unemploym	12.07
HCexpendpc	7.00
FemUnempl	11.52
FemSchool	53.08
LifeExpect	3.54
DPT	8.82
Measles	9.17
Population	2.93

Table 7: Variance Inflation Factors - Table 2

	vif
GDPpc	10.87
Rural	2.24
CO2	3.09
HCexpend	1.59
Water	3.81
Sanitation	4.34
HCexpendpc	6.21
FemUnempl	1.43
FemSchool	1.36
LifeExpect	3.06
DPT	8.59
Measles	8.89
Population	1.19

Data Imputation Matrix





Dahlgren, Göran, and Margaret Whitehead. 1991. “Policies and Strategies to Promote Social Equity in Health.” *Institute for Futures Studies*.

Haacker, Markus. 2002. “The Economic Consequences of HIV/AIDS in Southern Africa.” *IMF Working Paper* 02 (38).

Hurrelmann, Klaus. 1989. “Human Development and Health.” *Springer-Verlag*.

Kalemli-Ozcan, Sebnem. 2011. “AIDS, "Reversal” of the Demographic Transition and Economic Development: Evidence from Africa.” *Springer-Verlag*.

Solar, O, and A Irwin. 2010. “A Conceptual Framework for Action on the Social Determinants of Health. Social Determinants of Health Discussion Paper.” *Policy and Practice*, no. 2.

Talbott, John R. 2007. “Size Matters: The Number of Prostitutes and the Global HIV/AIDS Pandemic.”

PLoS ONE.

United Nations. 2014. "Millennium Development Goals." http://www.un.org/millenniumgoals/aids.shtml.

World Bank. 2014. http://microdata.worldbank.org/index.php/catalog/dhs/about.

World Health Organization, and UNICEF. 2014. "Joint Monitoring Programme for Water Supply and Sanitation." http://www.wssinfo.org/definitions-methods/data-sources/.