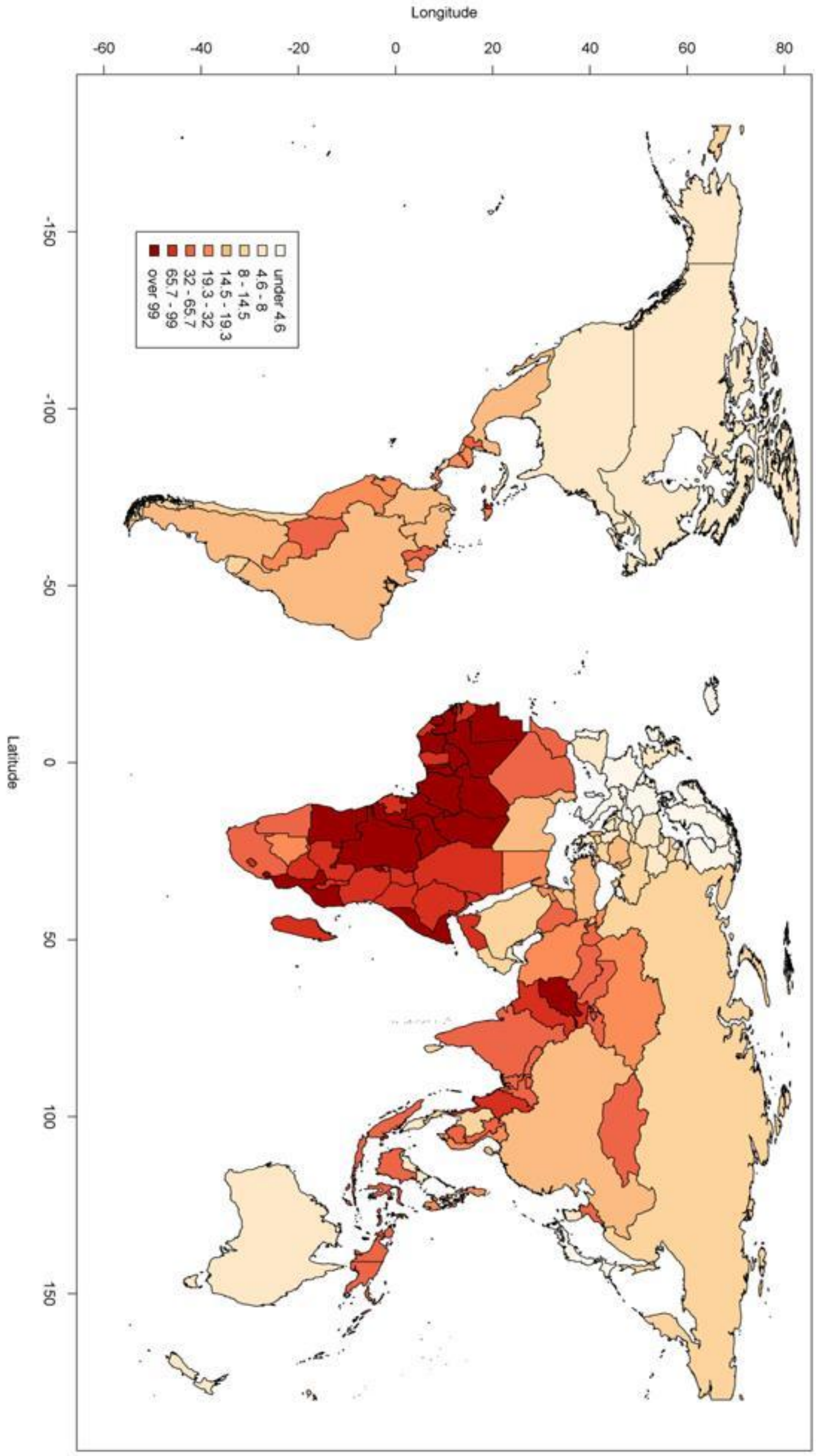
Appendix 1: World Bank estimations of U5MR (Under-five children death per 1000 live births) in 2009



Appendix 2: Multiple Imputation

MI was performed with the Amelia II R package [1]. Amelia II uses a bootstrapping-based EM algorithm that is both fast and robust [2]. Amelia II package was preferred to other available software due to its extensive use in social and economic sciences and its particular ability to deal well with missing data in longitudinal studies.

Description of the missing data pattern

Table 2 gives the percentage of missing data by studied variables between 1996 and 2009. The 5 variables with the largest percentage of missing data were the HIV prevalence (20.2%), the adolescent fertility rate (9.3%), the prevalence of undernourishment in the population (8.3%), the female mean years in school (7.2%), and the percentage of the population having access to improved sanitation facilities (5.3%). All other explanatory variables had a missing data percentage inferior to 5%.

Table 3 gives the percentage of missing data by country between 1996 and 2009. Countries with more than 30% of missing data are colored in red. These countries represented less than 5% of the whole countries in our dataset. The five countries with the most important missing data percentage were the South Sudan (86.7%), Andorra (64.8%), San Marino (63.3%), the Liechtenstein (55.2%), and Monaco (50%). In spite of this high level of missing data for these countries, we choose in our study to include them whatever their percentage of missing data.

Table 4 provides the percentage of missing data by year between 1996 and 2009. The most important percentage of missing data were observed for the year 1996 (17.5%). All other years had a percentage inferior to 10%.

Table 5 provides the number and percentages of countries with totally unobserved data by studied variables between 1996 and 2009. The 5 variables with the largest percentage of countries with totally unobserved data were the HIV prevalence (24%), the female mean years in school (9%), the prevalence of undernourishment (9%), the adolescent fertility rate (7%) and the GDP per capita (6%). All other variables had a percentage inferior to 5%.

Imputation characteristics

Further supplementary variables not used in the analysis were added to the dataset in view to improve the accuracy of the imputations. These last were chosen due to their potential strong correlations with some incomplete variables. These supplementary variables were: The fertility rate [3], the tuberculosis incidence [3], the percentage of the population under 14 years of age [3], the political IV democracy index (Polity IV) [4], the life expectancy [3] and the ratio of female to male primary school enrolment [3]

Figure 2 shows that our dataset included variables failing to fit a multivariate normal distribution assumed by Amelia II to impute data. U5MR, the GDP per capita, the public health expenditure per capita and the tuberculosis incidence were log-transformed to normalize their distribution and avoid imputations depending too heavily on extreme data points. In order to make the distributions symmetric and unbounded, a logistic transformation was used for water and sanitation coverages, and for the percentage of the population in urban area. A square root transformation was used for the HIV prevalence, the prevalence of undernourishment in the population, the adolescent fertility rate and the fertility rate. Geographic area was included as a six level nominal variable while the World Bank income country group (four levels) and the polity IV democracy index were treated as ordinal variables. To improve the imputation accuracy, Amelia II permits to include lags and leads of variables into the imputation model. In our study, this was done for all the variables. The time was also taken into account with the help of a second-order polynomial function of year. To take into account logical bounds of variable not handled by previous transformations, Amelia can take draws from a truncated normal distribution in order to achieve imputations that satisfy these bounds. The bounds used in our study are given in Table 6.

Inspection of the plausibility of the imputations

A way of assessing the fit of the imputation model is overimputing i.e. sequentially treating each of the observed values as missing. For each observed value in turn several hundred imputed values of that observed value are generated. Amelia II overimputation diagnostic, runs this procedure through all of the observed values for a particular variable. The estimates of each observation can be plotted against the true values of the observation. On this graph, a y = x line indicates the line of perfect agreement. For each observation, 90% confidence intervals can also be plotted that allow to visually inspect the behavior of the imputation model. The overimputation diagnostics for each variable are plotted on Figure 3.

A very good agreement between the observed and the imputed values was found for the GDP per capita, the water and sanitation coverage, the adolescent fertility rate, the HIV prevalence, the prevalence of undernourishment, the public health expenditure, the percentage of the population living in urban areas and the female mean years in school. A less good agreement was however found for the perceived level of corruption, democracy and violence. According to the low level of missing data for these 3 variables (inferior to 2%), it was argued that this less good agreement was not prejudicial to the overall quality of the imputed datasets.

For HIV prevalence, because of the large within country correlations and the number of countries with totally unobserved data, the previous inspection could not be as pertinent as for the others variables to assess the overall validity of the imputation process. However, the use of tuberculosis incidence as a supplementary variable in the imputation process (known to be strongly correlated to HIV prevalence) might have limited the presence of important incoherencies in estimations. Furthermore, the results of the analysis performed on countries with complete data presented in appendix 6 confirmed that the identified relation between log(U5MR) and the HIV prevalence was similar whatever the data used and gave a supplementary argument for the plausibility of the imputations of missing data in HIV prevalence.

Bibliography

1 Honaker J, King G, Blackwell M. Amelia II: A Program for Missing Data. Cambridge, MA: : Harvard Univ. 2007.

2 Horton NJ, Kleinman KP. Much Ado About Nothing. *The American Statistician* 2007;**61**:79–90.

3 World Bank. The World Bank Database. 2011.http://data.worldbank.org/ (accessed 10 Jan2012).

4 Marshall, M., Jaggers K. Political Regime Characteristics and Transitions, 1800-2011. 2011.http://www.systemicpeace.org/polity/polity4.htm (accessed 1 Apr2013).

Table 2: Percentage of missing data by variables between 1996 and 2009.

Variable	% missing data
Year	0,0
Under-5 mortality	0,0
Sanitation coverage (%)	5.3
Water coverage (%)	4.8
GDP per capita	4.2
HIV prevalence (%)	21.2
Female mean years in school	7.2
Perceived violence	1.6
Perceived democracy	0.4
Perceived corruption	1.7
Public health expenditure per capita	2.1
% Population in urban area	0,0
% Undernourishment in population	8.3
Adolescent fertility rate	9.3

Table 3: Percentage of missing data by country

Country	% of missing data
Afghanistan	11
Albania	7.1
Algeria	7.1
Andorra	**64.8**
Angola	0.5
Antigua and Barbuda	**30.5**
Argentina	2.4
Armenia	0.5
Australia	0.5
Austria	0.5
Azerbaijan	0.5
Bahamas. The	2.9
Bahrain	27.1
Bangladesh	0.5
Barbados	7.1
Belarus	0.5
Belgium	2.4
Belize	0.5
Benin	0.5
Bhutan	14.8
Bolivia	0.5
Bosnia and Herzegovina	7.1
Botswana	0.5
Brazil	0.5
Brunei Darussalam	27.1
Bulgaria	0.5
Burkina Faso	0.5
Burundi	0.5
Cambodia	0.5
Cameroon	0.5
Canada	0.5
Cape Verde	1.4
Central African Republic	0.5
Chad	0.5
Chile	0.5
China	7.1
Colombia	0.5
Comoros	0.5
Congo. Dem. Rep.	13.8
Congo. Rep.	1.4
Costa Rica	0.5
Cote d Ivoire	0.5
Croatia	0.5
Cuba	7.1
Cyprus	7.1
Czech Republic	0.5
Denmark	0.5
Djibouti	0.5
Dominica	**33.8**
Dominican Republic	0.5
Ecuador	0.5
Egypt. Arab Rep.	0.5
El Salvador	0.5
Equatorial Guinea	10
Eritrea	1.4
Estonia	0.5
Ethiopia	0.5
Fiji	0.5
Finland	0.5
France	0.5
Gabon	0.5
Gambia. The	0.5
Georgia	0.5
Germany	0.5
Ghana	0.5
Greece	0.5
Grenada	16.2
Guatemala	0.5
Guinea	0.5
Guinea-Bissau	0.5
Guyana	0.5
Haiti	0.5
Honduras	7.1
Hungary	0.5
Iceland	7.1
India	7.1
Indonesia	0.5
Iran. Islamic Rep.	0.5
Iraq	14.3
Ireland	0.5
Israel	0.5
Italy	7.1
Jamaica	6.7
Japan	0.5
Jordan	7.1
Kazakhstan	0.5
Kenya	0.5
Kiribati	**32.4**
Korea. Dem. Rep.	27.1
Korea. Rep.	0.5
Kuwait	7.1
Kyrgyz Republic	0.5
Lao PDR	0.5
Latvia	1
Lebanon	2.4
Lesotho	0.5
Liberia	1.4

Country	% of missing data
Libya	19
Liechtenstein	**55.2**
Lithuania	1.4
Luxembourg	2.4
Macedonia. FYR	9
Madagascar	0.5
Malawi	0.5
Malaysia	0.5
Maldives	0.5
Mali	0.5
Malta	7.1
Marshall Islands	**42.4**
Mauritania	0.5
Mauritius	0.5
Mexico	0.5
Micronesia. Fed. Sts.	27.6
Moldova	0.5
Monaco	**50**
Mongolia	0.5
Montenegro	29
Morocco	0.5
Mozambique	0.5
Myanmar	13.8
Namibia	0.5
Nepal	0.5
Netherlands	0.5
New Zealand	7.1
Nicaragua	0.5
Niger	0.5
Nigeria	0.5
Norway	0.5
Oman	13.8
Pakistan	0.5
Palau	**48.6**
Panama	0.5
Papua New Guinea	7.1
Paraguay	0.5
Peru	0.5
Philippines	0.5
Poland	7.6
Portugal	0.5
Qatar	9
Romania	2.4
Russian Federation	7.1
Rwanda	0.5
Samoa	13.8
San Marino	**63.3**
Sao Tome and Principe	2.9
Saudi Arabia	18.6
Senegal	0.5
Serbia	10
Seychelles	**33.8**
Sierra Leone	0.5
Singapore	7.6
Slovak Republic	0.5
Slovenia	0.5
Solomon Islands	11.9
Somalia	20.5
South Africa	0.5
South Sudan	**86.7**
Spain	0.5
Sri Lanka	0.5
St. Kitts and Nevis	**36.7**
St. Lucia	14.3
St. Vincent and the Grenadines	28.1
Sudan	6.7
Suriname	0.5
Swaziland	0.5
Sweden	0.5
Switzerland	0.5
Syrian Arab Republic	7.1
Tajikistan	0.5
Tanzania	0.5
Thailand	0.5
Timor-Leste	14.8
Togo	0.5
Tonga	18.1
Trinidad and Tobago	0.5
Tunisia	0.5
Turkey	0.5
Turkmenistan	8.6
Tuvalu	**51.9**
Uganda	0.5
Ukraine	0.5
United Arab Emirates	7.1
United Kingdom	0.5
United States	0.5
Uruguay	0.5
Uzbekistan	7.1
Vanuatu	8.1
Venezuela. RB	2.4
Vietnam	0.5
West Bank and Gaza	17.6
Yemen. Rep.	0.5
Zambia	0.5
Zimbabwe	11

Table 4: Percentage of missing data by year.

Year	% missing data
1996	17.5
1997	9.1
1998	8.7
1999	8.5
2000	8.1
2001	8.1
2002	7.9
2003	7.6
2004	7.7
2005	7.6
2006	7.8
2007	8.1
2008	8.2
2009	8.5

Table 5: Number (percentage) of countries with totally unobserved data during 2000-2009 by variables

Variables	Number of countries (%)
Sanitation	10 (5%)
Water	8 (4%)
GDP per capita	12 (6%)
HIV prev.	47 (24%)
Female school	18 (9%)
Violence	2 (1%)
Democracy	2 (1%)
Corruption	5 (3%)
Public health exp.	5 (3%)
Urbanization	0 (0%)
Undernourishment	18 (9%)
Ado. Fertility rate	13 (7%)

Table 6: Variable bounds used in the imputation process

Variable	Minimum value	Maximum value
GDP per capita	0	100000
HIV prevalence (%)	0	40
Female mean years in school	0	20
Perceived violence	-3,5	3
Perceived democracy	-3	3
Perceived corruption	-3	3
Public health expenditure per capita	0	10000
Adolescent fertility rate	0	300
Fertility rate	0	15
Under-14 population	5	60
Democracy (Polity IV)	-10	10
Tuberculosis incidence	0	2500
Life expectancy	20	100
Ratio of Female to male primary enrolment	0	150

Figure 2: Observed distributions of the study variables

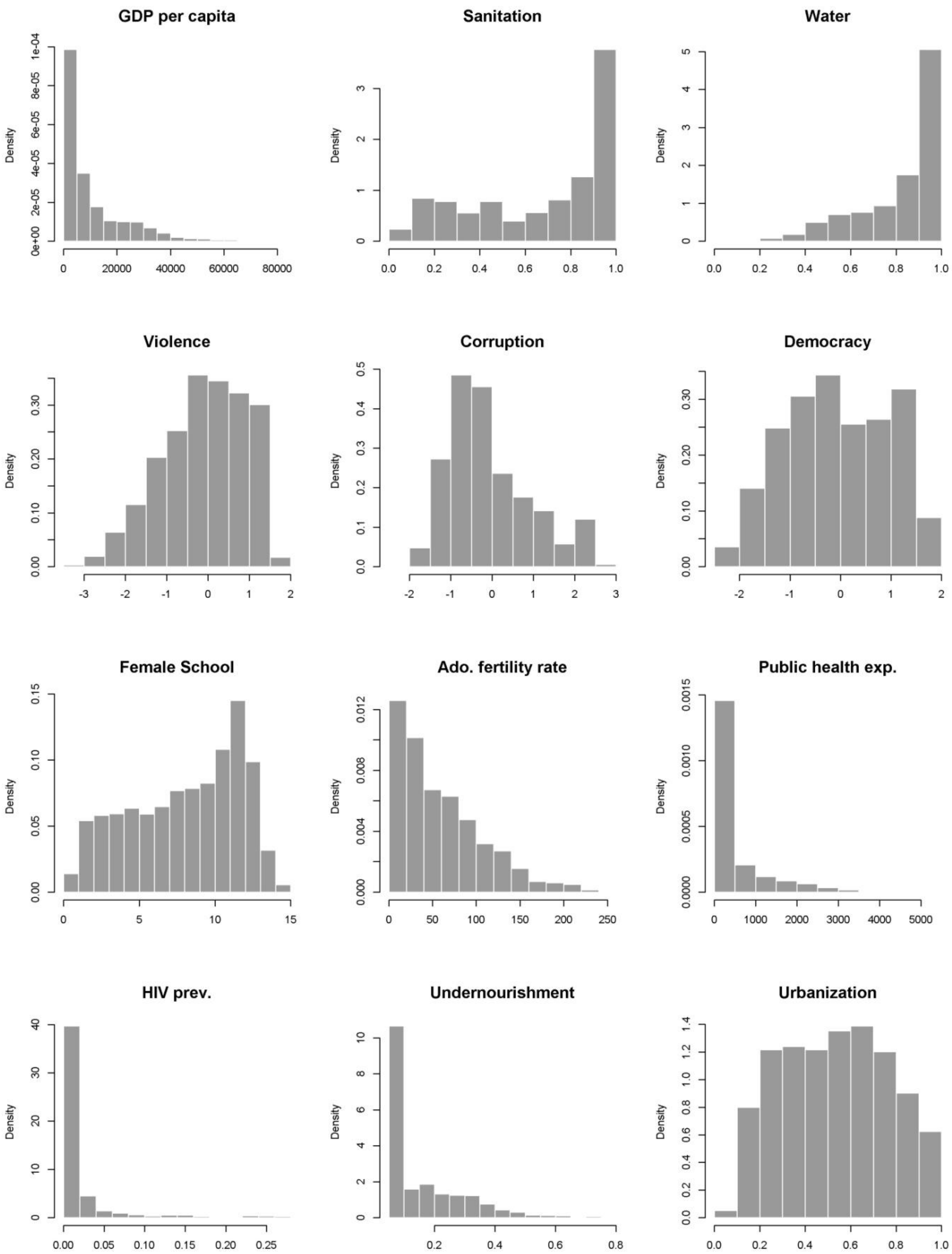
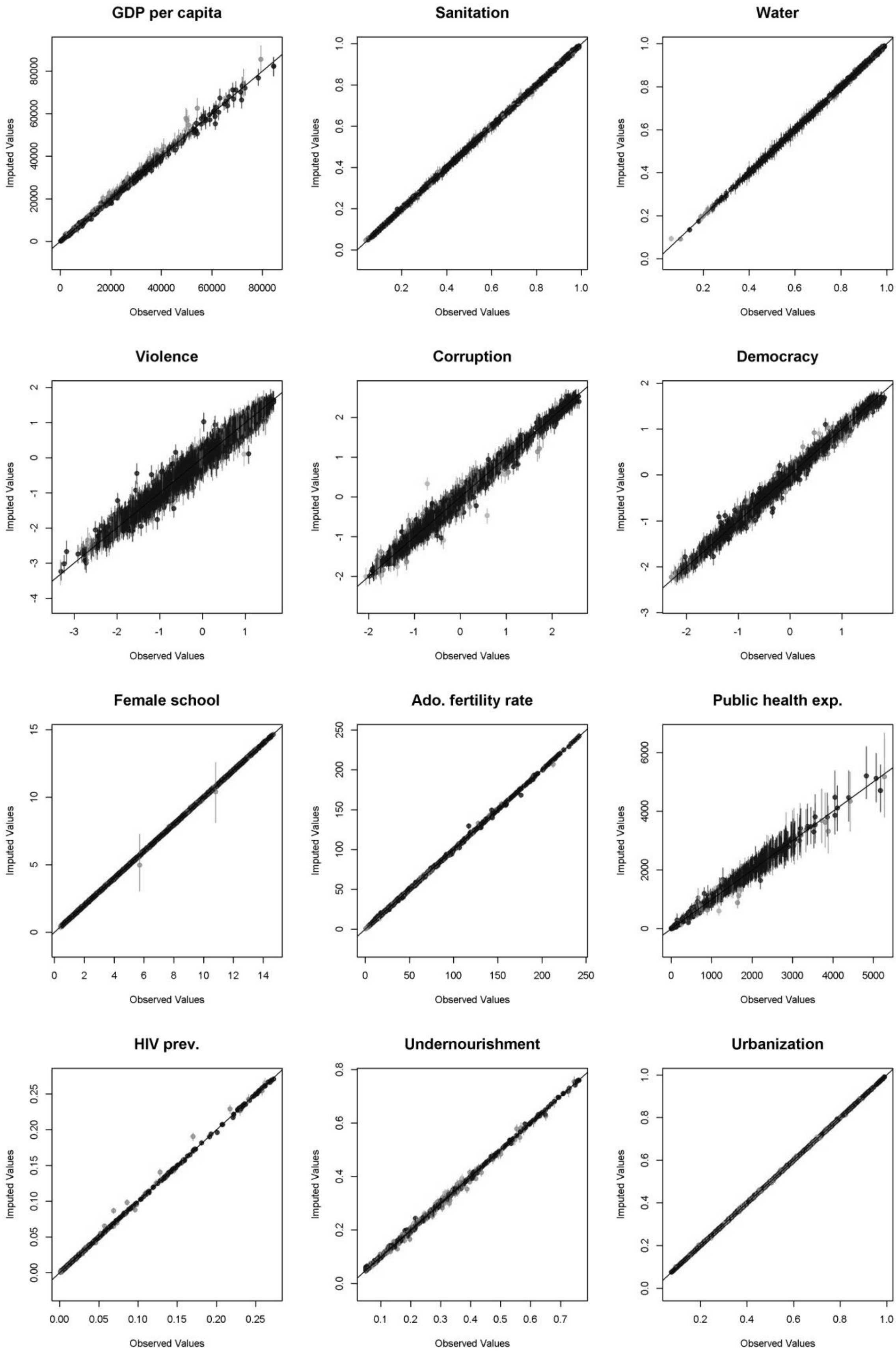


Figure 3: Overimputation diagnostic graphs for the studied variables



Appendix 3: Stepwise procedure

Variable selection methodology

A methodological problem associated with variable selection for the final model arose from the MI procedure. Here, as suggested in a previous study, we performed our variable selection procedure on each of the imputed datasets created by multiple imputation and selected predictors that appeared in more than half of the models.

To avoid over-adjustment and problems in algorithm convergence due to the large numbers of exploratory variables (36 in all) we chose to use a forward stepwise procedure. New variables were added manually one at a time and the variable that minimized the Akaike Information Criterion (AIC) was retained:

- Step 1: A single variable GAMMs was run for each variable and they were ranked in order of AIC. The variable with the lowest AIC was chosen as the first variable for the model.
- Step 2: Each of the remaining variables was added in turn to the one variable model from Step 1 and the AIC re-calculated. The additional variable that produced the lowest AIC was then selected.
- Step 3: Step 2 was repeated. The process was repeated until no additional variable decreased the AIC.

Result of stepwise procedures

Selected variables in the 10 stepwise procedures on each of the imputed dataset are given in Table 7. According to our decision rule the following variables were retained in the final model:

- The GDP per capita without lag
- The percentage of people with access to improve water source with a lag of 4 years
- The percentage of people with access to improved sanitation with lags of 0 and 4 years
- The percentage of urban population with a lag of 4 years
- The adolescent fertility rate with lags of 0 and 4 years
- The public health expenditure with lags of 0, 2 and 4 years
- The HIV prevalence without lag
- The corruption level with lags of 0 and 4 years
- The political stability level with a lag of 4 years
- The average years of schooling for women with lags of 2 and 4 years

Globally there was a strong agreement between the 10 analyses as for the name of the selected variables but fluctuations in the selection of some lags. It can be noted that GDP at lag 0 was selected in each of the 10 imputed datasets but that GDP at lag 2 was also selected in 5 datasets and GDP at lag 4 was selected in 4 other datasets. Thus, some form of delayed effect of GDP is likely. By contrast, democracy was never selected, whatever the lag, in any of the 10 datasets and prevalence of undernourishment was only selected once (lag 2 and 4 simultaneously).

Table 7: Summary of the stepwise procedures realized on the 10 datasets produced by MI (variables in blue were those selected by the associated stepwise procedure and variables in pink those that were not selected).

Variables	Lag	Dataset 1	Dataset 2	Dataset 3	Dataset 4	Dataset 5	Dataset 6	Dataset 7	Dataset 8	Dataset 9	Dataset 10
Demographic and socio economic factors											
GDP per capita	*Lag 0**										
	Lag 2										
	Lag 4										
People with access to improved water source	*Lag 0*										
	Lag 2										
	*Lag 4**										
People with access to improved sanitation	*Lag 0**										
	Lag 2										
	*Lag 4**										
Percentage of urban population	*Lag 0*										
	Lag 2										
	*Lag 4**										
Adolescent fertility rate	*Lag 0**										
	Lag 2										
	*Lag 4**										
Health and medical factors											
Public health expenditure per capita	*Lag 0**										
	*Lag 2**										
	*Lag 4**										
HIV prevalence (%)	*Lag 0**										
	Lag 2										
	Lag 4										
Prevalence of undernourishment	*Lag 0*										
	Lag 2										
	Lag 4										
Political and societal factors											
Corruption level	*Lag 0**										
	Lag 2										
	*Lag 4**										
Democracy	*Lag 0*										
	Lag 2										
	Lag 4										
Political Stability	*Lag 0*										
	Lag 2										
	*Lag 4**										
Average years of schooling for women	*Lag 0*										
	*Lag 2**										
	*Lag 4**										

* Variables selected in the final model (variables that appear in more than half of the models)

Appendix 4: Stability of primary results - stratification

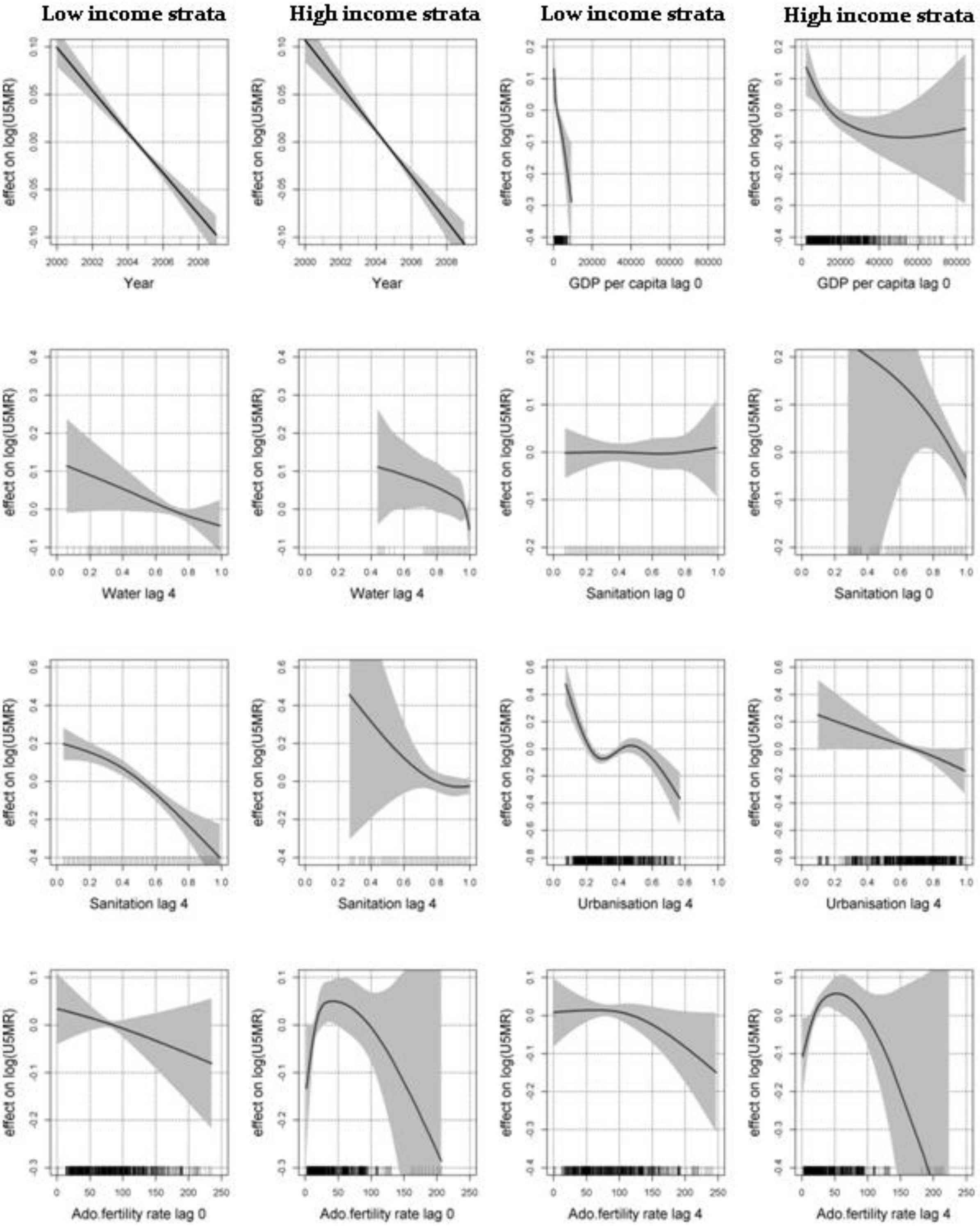
Stratified analyses according to national income were conducted to examine the stability of the primary results. The 2011 country income groups defined by the World Bank were used. To limit the loss of power associated with small sample size, only two strata were considered: in the low income strata, low income countries (LIC) were merged with lower middle income countries (LMC) and in the high income strata, upper middle income countries (UMC) were merged with high income countries (HIC). The previously identified final model was run separately on each associated subset. Response plots from the two final GAMM are shown in the Figures 4 and 5.

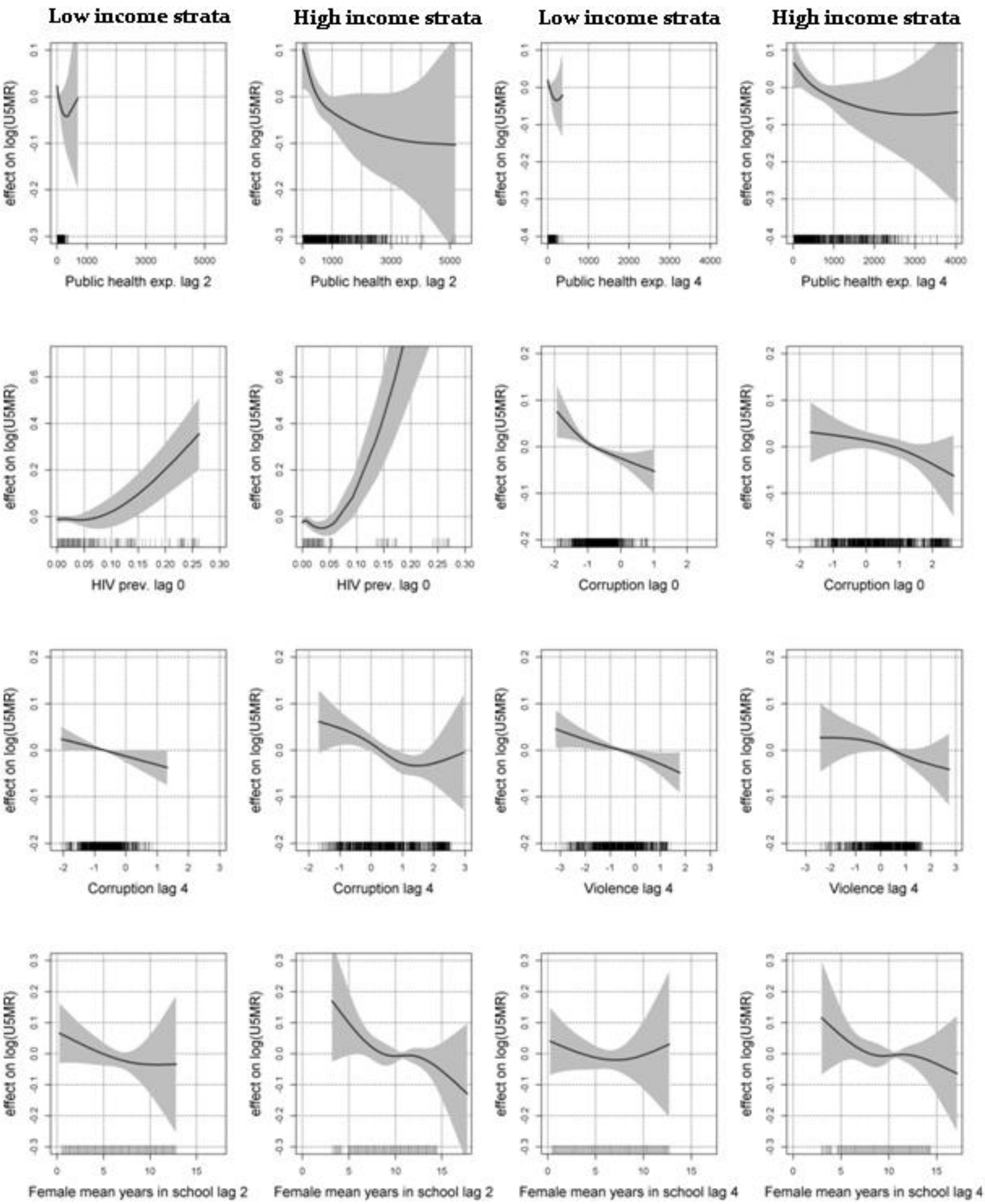
The same functional forms of the relation with U5MR were found for the time period, the GDP per capita at lag 0, the percentage of people with access to improved sanitation facilities at lag 4, the HIV prevalence at lag 0, the perceived level of corruption at lags 0 and 4, the percentage of the population with access to improved water sources with at lag 4, the proportion of the population living in urban areas with at lag 4, the public health expenditure at lags 2 and 4, the perceived level of violence with at lag 4 and the female mean years in school at lags 2 and 4.

Clear discordances in the functional form with U5MR were found for the proportion of the population having access to improved sanitation facilities at lag 0 and for the adolescent fertility rate at lags 0 and 4. These variables were only significant for higher income countries.

These results validated the overall stability of the primary results. Discordances found in the stratified analyses were however discussed in details in the main part of the article.

Figure 4: Relations between the logarithm of U5MR in 1) the countries with a lower income (LIC and LMC) and the countries with a higher income (UMC and HIC) and 2) the sixteen continuous variables included in the final GAMM model. The y-axis is the effect of the variable; grey areas are the 95% confidence intervals. Rug plots on the x-axis are the observed values.




Low income strata
High income strata
Low income strata
High income strata
effect on log(U5MR)
Public health exp. lag 2
Public health exp. lag 2
Public health exp. lag 4
Public health exp. lag 4
HIV prev. lag 0
HIV prev. lag 0
Corruption lag 0
Corruption lag 0
Corruption lag 4
Corruption lag 4
Violence lag 4
Violence lag 4
Female mean years in school lag 2
Female mean years in school lag 2
Female mean years in school lag 4
Female mean years in school lag 4

Appendix 5: Stability of results – Use of alternative models specifications

To assess the stability of the results, alternative model specifications were used. The previously identified final model was run with an alternative number of knots (k=6) and an alternative smoothing basis (thin plate regression spline). Response plots from these two alternative GAMM are shown in Figures 6 and 7. No differences with the original model were found confirming the overall stability of the final GAMM model results.

Figure 6: Relations between the logarithm of U5MR and the sixteen continuous variables included in the final GAMM model (obtained with the use of a thin plate regression spline as the smoothing basis). The y-axis is the effect of the variable; grey areas are the 95% confidence intervals. Rug plots on the x-axis are the observed values.

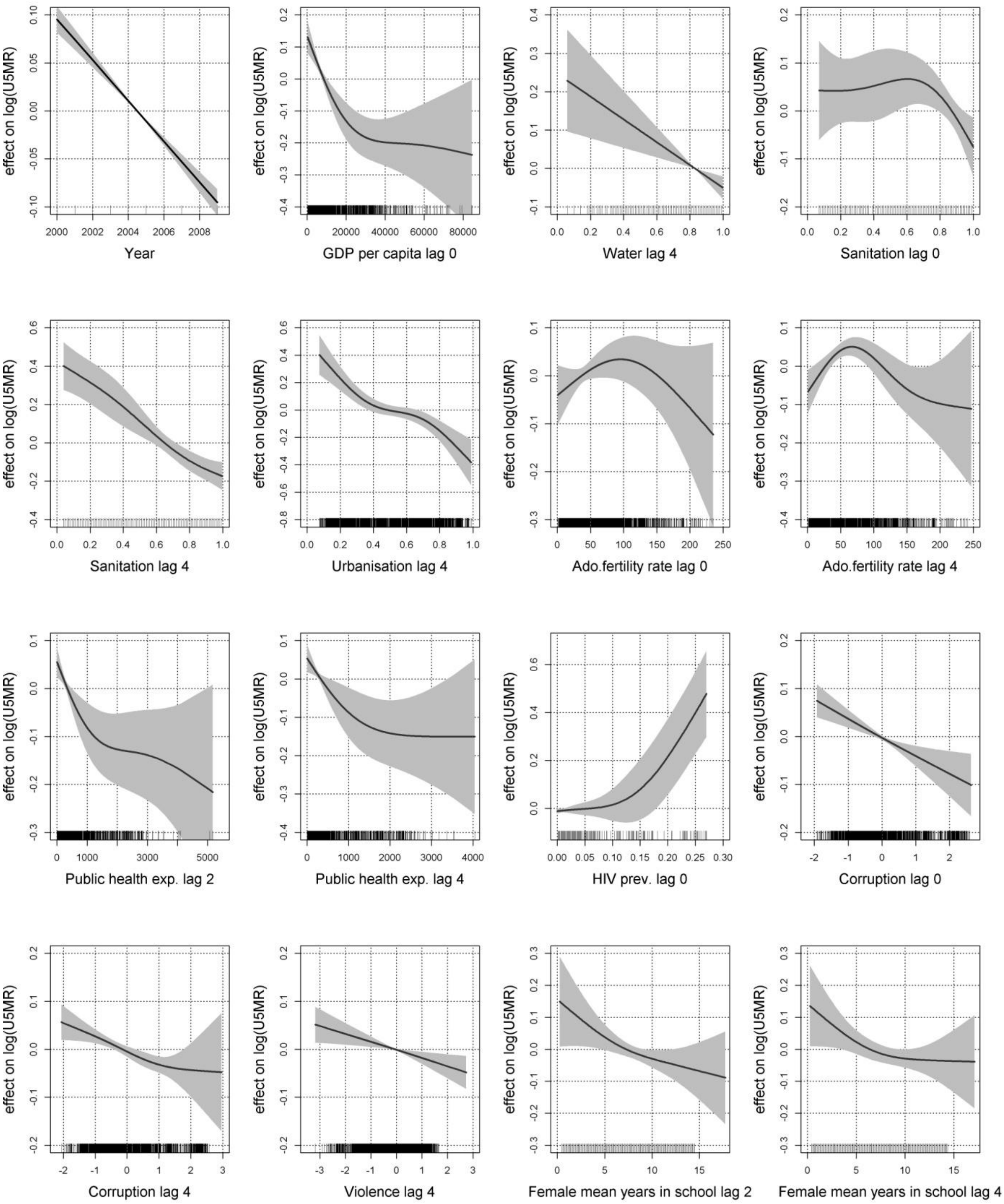
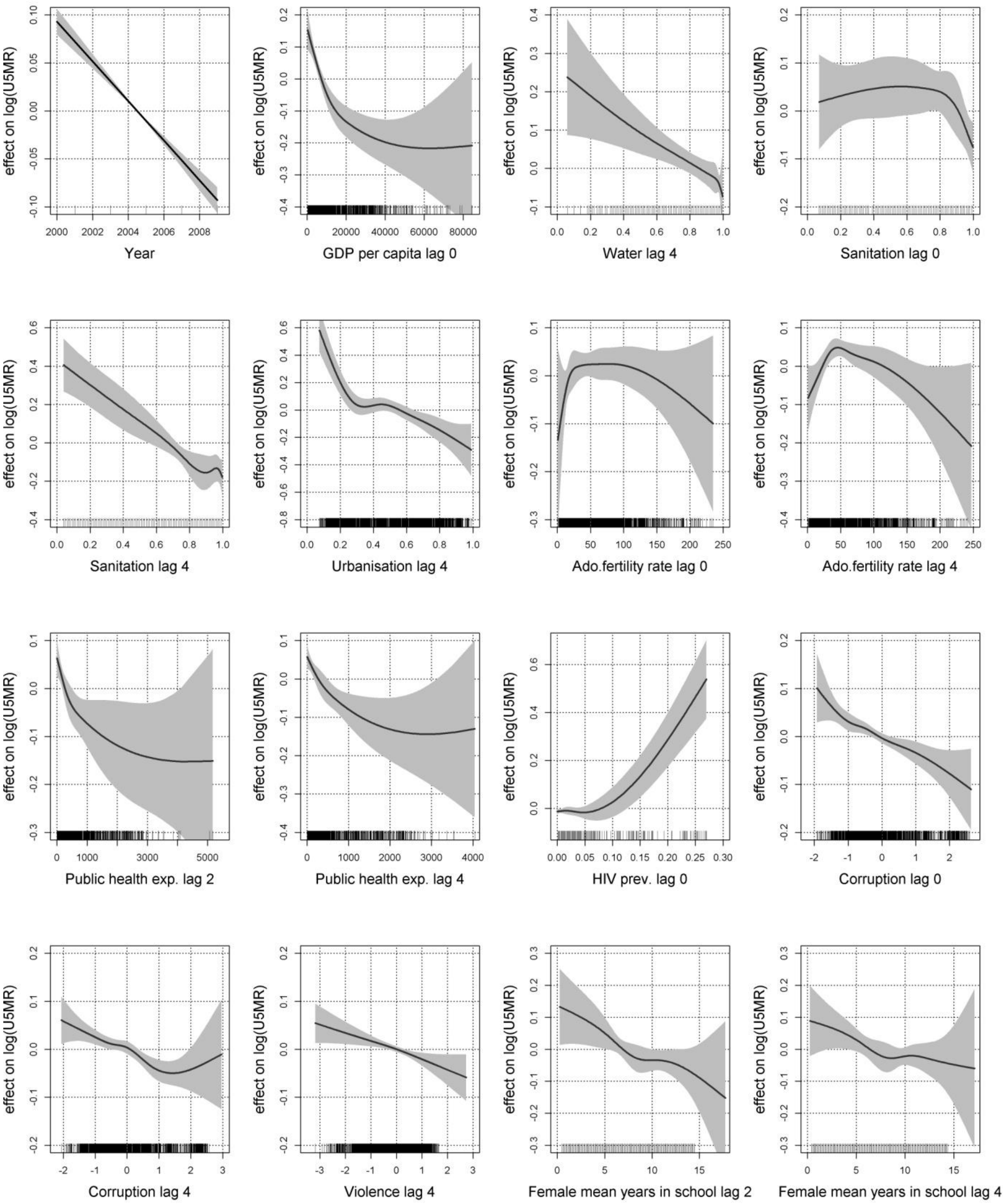


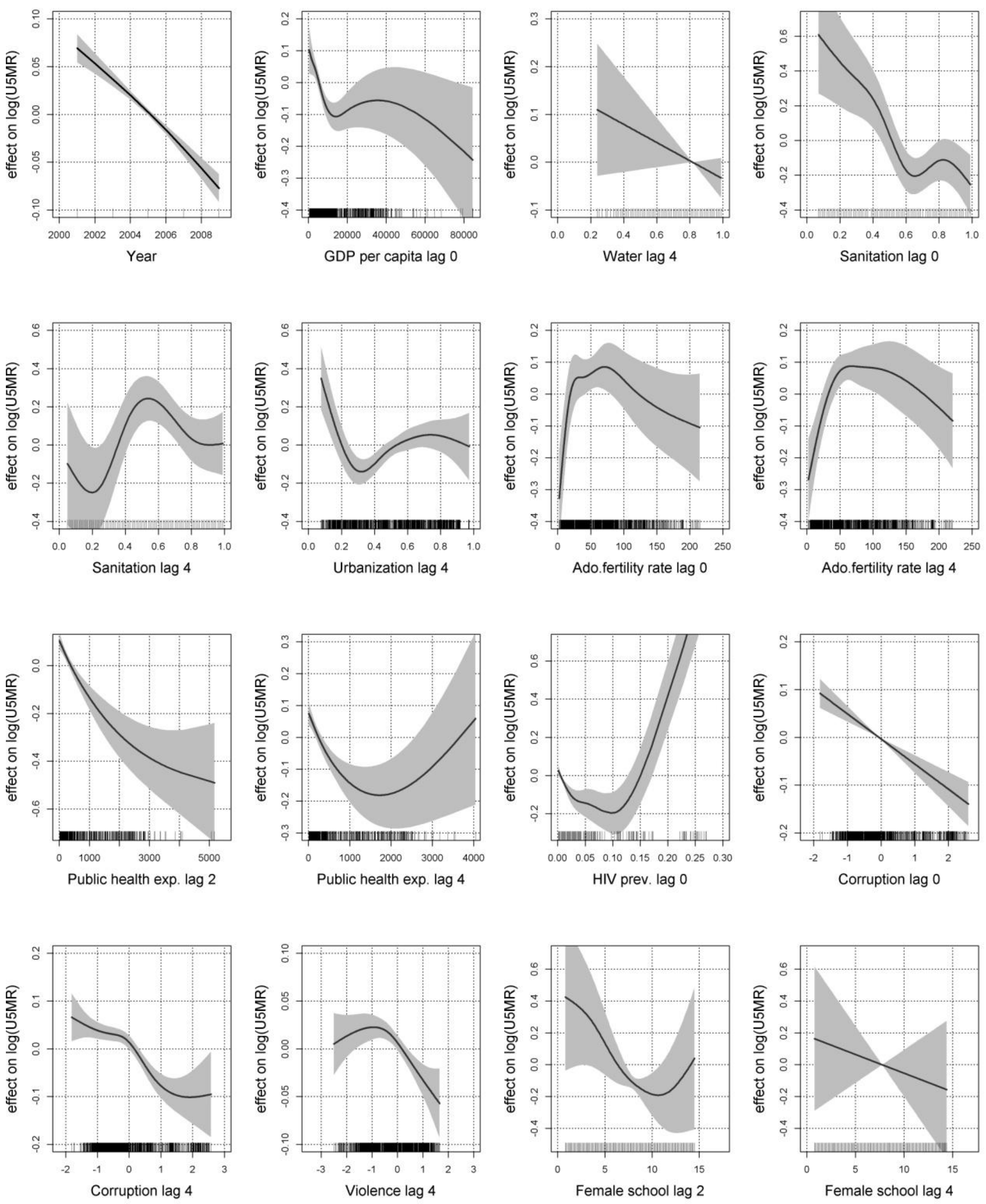
Figure 7: Relations between the logarithm of U5MR and the sixteen continuous variables included in the final GAMM model (obtained with an alternative number of knots k=6). The y-axis is the effect of the variable; grey areas are the 95% confidence intervals. Rug plots on the x-axis are the observed values.



Appendix 6: Stability of results – Analysis of completely observed countries

To assess the stability of the primary results, the previously identified final model was run on the subset of completely observed countries. Response plots are shown in Figure 8.

Figure 8: Relations between the logarithm of U5MR and the sixteen continuous variables included in the final GAMM model obtained with complete data only. The y-axis is the effect of the variable; grey areas are the 95% confidence intervals. Rug plots on the x-axis are the observed values.

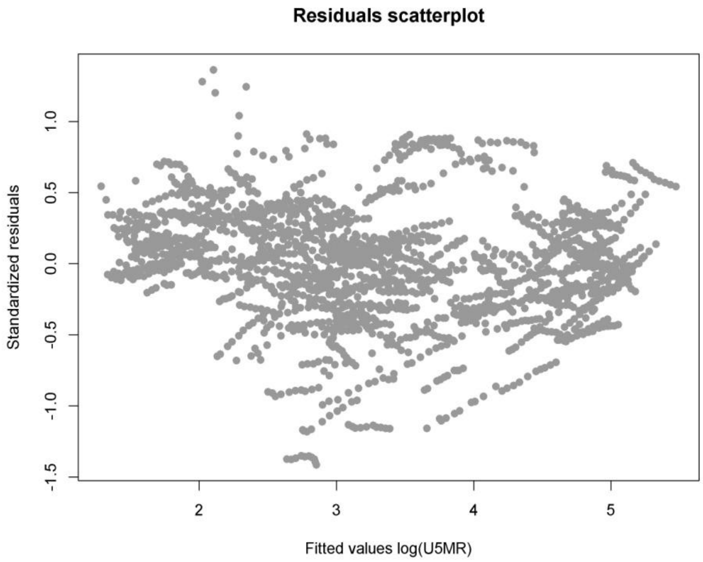


Appendix 7: Validity of the final GAMM Model

To check the validity of the final GAMM model, the homoscedasticity of residuals was assessed graphically (Figure 9). The distribution of the observed standardized residuals did not reveal any value abnormally small or large (Figure 9). The qqplot mostly agreed with the normality assumption (Figure 10). It can be however noted that the largest negative standardized residuals (between -1.0 and -1.5) came mainly from the central Asian part of the Europe-central Asia region. The heterogeneity between the two components of this region is also responsible for the imperfect fit of the model reflected in the bottom-left part of the qqplot. Sensitivity analyses separating central Asia from Europe region did not alter at all the retrieved relations between determinants and U5MR.

Absence of residual correlations was also checked graphically by income groups and year of sampling (Figure 11). Although residuals were a little higher for HIC compared to other countries, graphics reveals no particular strong pattern of correlation confirming the overall validity of the final GAMM model.

Figure 9: Residuals scatterplot and predicted versus observed values scatterplot in the final GAMM model



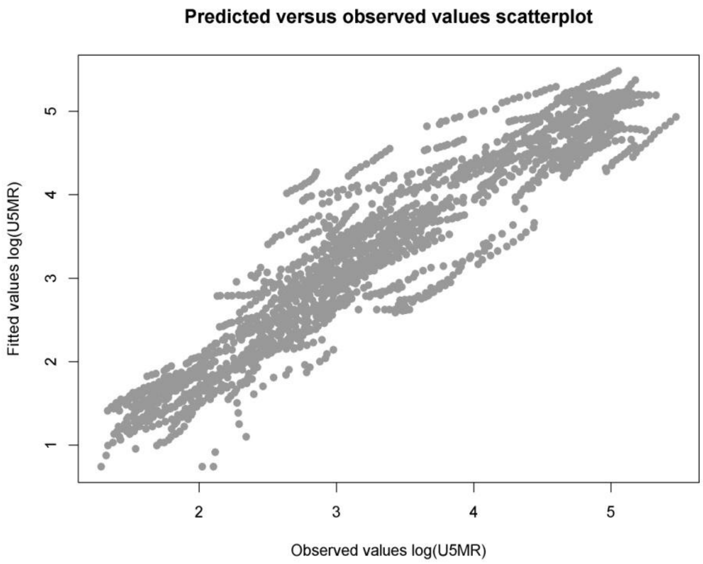
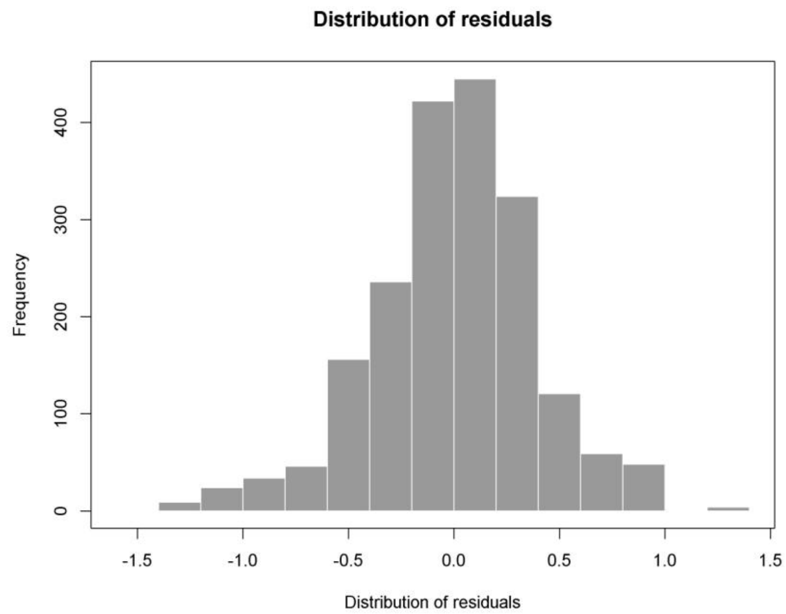


Figure 10: Distribution of the residuals and Q-Q plot in the final GAMM model



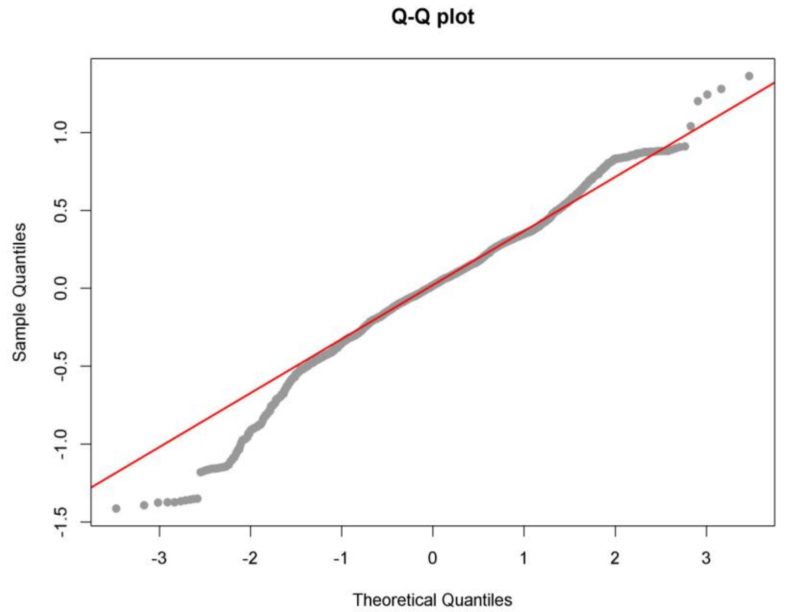
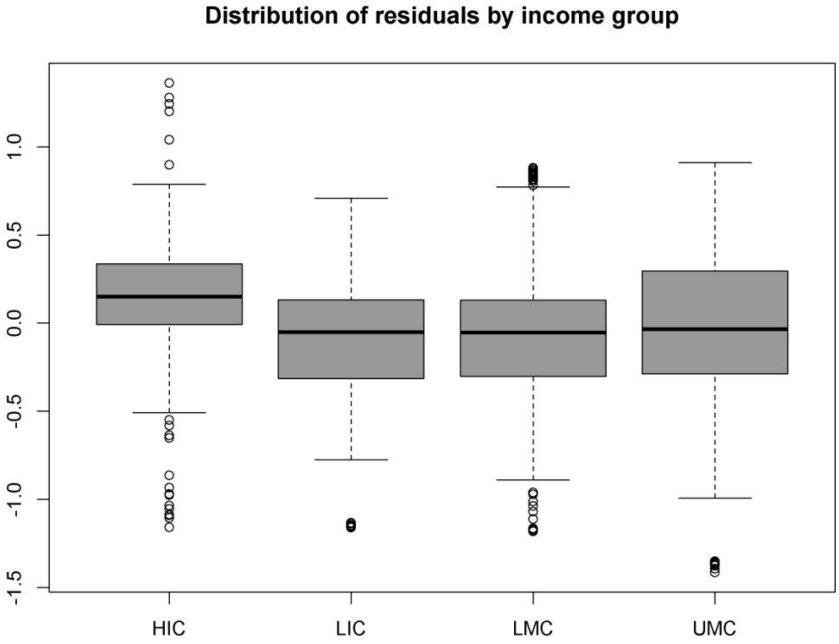
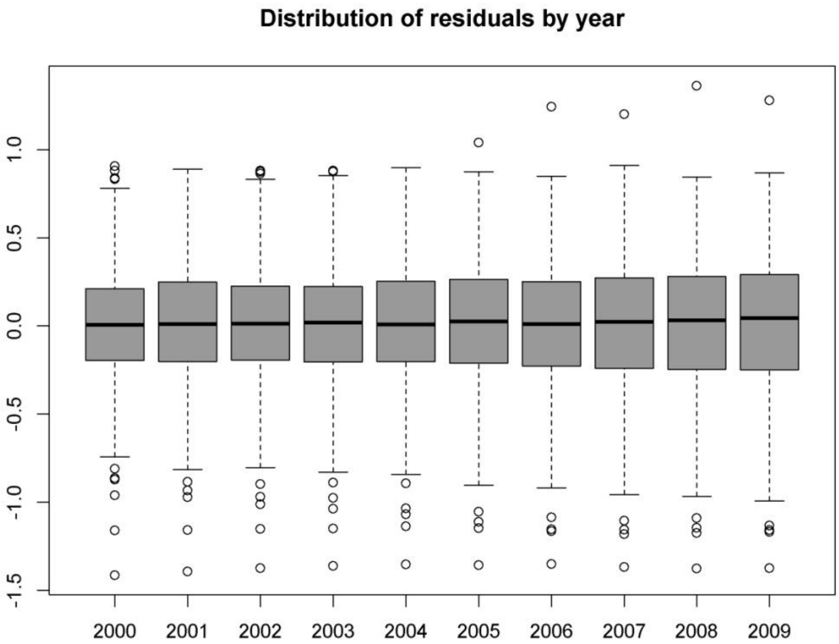


Figure 11: Distributions of residuals in the final GAMM model by income group and year of sampling.





*HIC: High Income Countries, LIC: Low Income Countries, LMC: Lower Middle income Countries, UMC: Upper Middle income Countries

Appendix 8: Correlations between the study variables

Crude Pearson's correlations between unlagged quantitative variables used in this analysis are presented in Table 8.

The 5 variables the most correlated with the logarithm of U5MR were: the percentage of the population having access to improved sanitation facilities (r = - 0.83), the female mean years in school (r = -0.82), the percentage of the population having access to improved water sources (r = -0.79), the perceived level of corruption (r = -0.76) and the GDP per capita (r = -0.75)

Table 8: Crude Pearson's correlations between quantitative variables used in this analysis.

	U5MR (log)	Sanitation	Water	GDP	HIV prev.	Female school	Violence	Democracy	Corruption	Health exp.	Urbanization	Under nutrition	Ado. Fert. rate
U5MR (log)	-	-	-	-	-	-	-	-	-	-	-	-	-
Sanitation	-0.83	-	-	-	-	-	-	-	-	-	-	-	-
Water	-0.79	0.79	-	-	-	-	-	-	-	-	-	-	-
GDP per capita	-0.75	0.59	0.52	-	-	-	-	-	-	-	-	-	-
HIV prev.	0.37	-0.32	-0.24	-0.19	-	-	-	-	-	-	-	-	-
Female school	-0.82	0.81	0.75	0.57	-0.17	-	-	-	-	-	-	-	-
Violence	-0.6	0.47	0.52	0.56	-0.05	0.53	-	-	-	-	-	-	-
Democracy	-0.64	0.42	0.51	0.49	-0.11	0.56	0.67	-	-	-	-	-	-
Corruption	-0.76	0.54	0.56	0.76	-0.12	0.58	0.72	0.76	-	-	-	-	-
Public Health exp.	-0.73	0.53	0.47	0.86	-0.17	0.56	0.54	0.63	0.79	-	-	-	-
Urbanization	-0.67	0.61	0.6	0.64	-0.26	0.57	0.38	0.38	0.55	0.55	-	-	-
Undernourishment	0.66	-0.66	-0.67	-0.47	0.29	-0.61	-0.5	-0.46	-0.48	-0.42	-0.57	-	-
Ado. Fertility rate	0.74	-0.72	-0.7	-0.51	0.32	-0.68	-0.44	-0.32	-0.49	-0.48	-0.47	0.5	-