COVID-19 in Africa

Econ 2509, Fall ‘21
Group 10: Sejoon P., Sneha B., Tyler W

Table of Contents

I. Abstract
II. Hypothesis
III. Data
IV. Exploratory Analysis
    A. Descriptive Statistics
    B. Plot 1: COVID Deaths Histogram
    C. Plot 2: COVID Deaths vs age65 Regression
V. Regression Analysis
    A. Baseline Regression
    B. Alternative Regression
    C. Hypothesis Tests
VI. Data Shortcomings
VII. Conclusion

I. Abstract

COVID-19 has spread worldwide with more than 260M cases reported, resulting in 5.2M deaths. COVID has impacted the world's most developed economies by stifling trade, travel, GDP, and a host of other economic indicators. Africa is the world's poorest continent as defined by people living in poverty, and therefore includes a population that is vulnerable to the spread of infectious disease. Africa is also the continent with the least amount of vaccinated individuals. Therefore, we collected and analyzed data on COVID-19 deaths for each country in Africa, and have chosen variables we believe to be indicators in driving deaths.

II. Hypothesis

COVID-19 deaths are likely to be higher in more populated, less developed countries.

Less developed: We use this term to mean a country with a less robust economy as it relates to education, development indicators (AIDI), lower healthcare costs per capita, and a country with a low political stability score.

III. Data

- Original dataset from Kaggle
  - Country, total cases, total deaths, total recovered, total population
  - Compiled from respective government websites
- Added data from the World Bank
  - 2019 data, or most recent available
- Self-calculated fields
  - logPop
  - deathPer
- Dummy variable
  - Dict
- **<u>Variable of interest: deaths</u>**

deaths	total # of COVID-19 deaths in a country
pop	population of country
logPop	log of population of country
deathPer	COVID-19 deaths as a % of country's population
educ	number of years of compulsory education
unemp	unemployment rate
corr	country's corruption level, from a scale of -2.5 to 2.5
polStab	country's political stability level, from a scale of -2.5 to 2.5
FDI	FDI as a % of country's GDP
AIDI	country's development index score, from a scale of 0 to 100
UHCI	country's essential health services index score, from a scale of 0 to 100
avgTemp	country's average year long temperature
dict	dummy variable indicating 1 if a country is a dictatorship and 0 otherwise
popDensity	population density of a country, measured in people per sq km of land
HEPC	healthcare expenditure per capita, measured in USD
age65	% of population that's age 65 years+
diabetes	% of population that has diabetes
handwashing	% of population that has access to handwashing facilities

IV. Exploratory Analysis: Descriptive Statistics

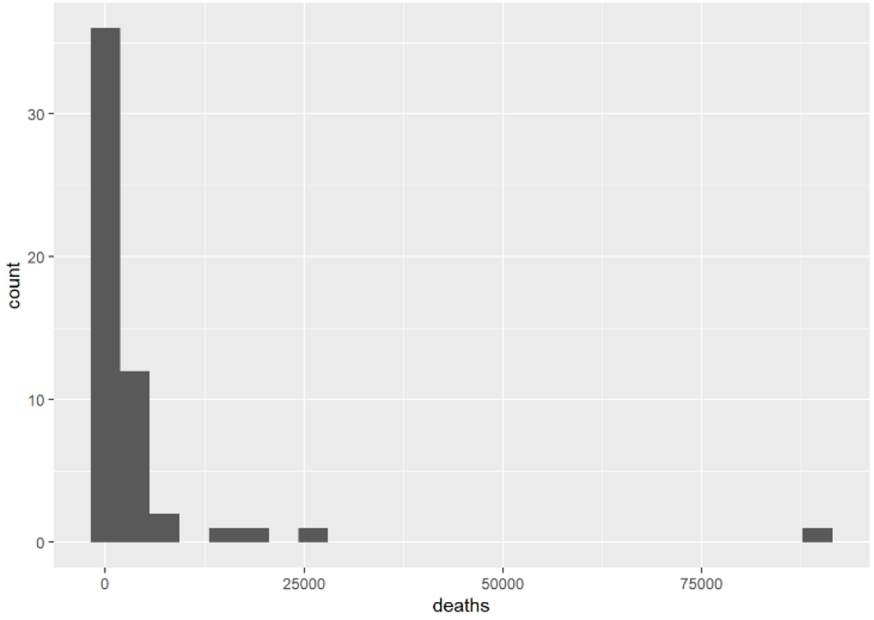
COVID-19 in Africa

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
COVID Deaths	54	4,099.54	12,751.40	38	238.5	770.5	2,833.8	89,562
Healthcare per cap	54	372.86	530.65	25.00	112.79	155.55	377.12	3,207.47
Population	54	25,612,298.00	36,983,951.00	99,189	2,856,880.0	13,509,553.0	31,144,276.0	213,187,117
Pop Density	54	117.30	149.10	3.09	25.59	56.16	130.05	623.52
Political Stability	54	-0.64	0.84	-2.52	-1.14	-0.45	-0.20	1.09
Mandatory Education	54	8.28	1.98	5	7	8	10	15

- There are 54 countries included in the dataset, which includes all countries in Africa
- The average healthcare spend per capita is $372.86 USD, compared to ~$10,600 in the USA
- The average years of compulsory schooling is 8 with a standard deviation of ~2 years.
- The average population size is 25.612M.
- Political stability is relatively low on the continent, sitting at -.64 on a scale from -2.5 to 2.5, with 2.5 being the most political stable a country can be.
- The average population density is 117 people per square mile; for context, the USA has a population density of ~93 people per square mile
- South Africa has, by far, the most deaths at 89,562 while Burundi has the least at 38

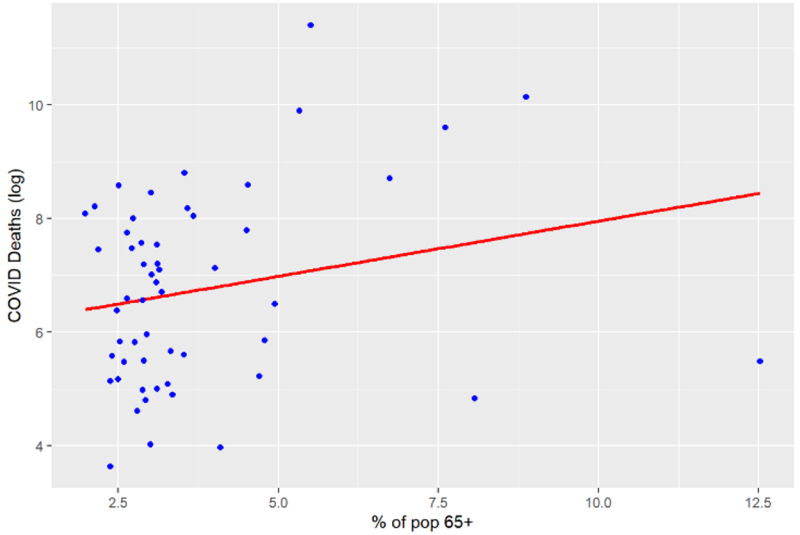
IV. Exploratory Analysis: Plots

Plot 1: Distribution of COVID-19 Deaths



- Confirms distribution of deaths is not normal, so we will transform using log of deaths

Plot 2: COVID-19 Deaths in Africa vs % of Population that's 65 and over



- Older countries face more COVID deaths

V. Regression Analysis: Baseline

COVID in Africa - Baseline Regressions

	Dependent variable:					
	log(deaths)					
	(1)	(2)	(3)	(4)	(5)	(6)
age65	0.174	0.209	0.328**	-0.030		
	(0.186)	(0.135)	(0.113)	(0.101)		
HEPC	0.0002	0.0004	0.001	0.001*	0.001*	0.001*
	(0.0004)	(0.0004)	(0.0004)	(0.0003)	(0.0003)	(0.0003)
popDensity		-0.004*	-0.003*	-0.003*	-0.003*	-0.003*
		(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
logPop			1.589***	1.648***	1.657***	1.669***
			(0.197)	(0.188)	(0.183)	(0.188)
AIDI				0.045***	0.043***	0.046***
				(0.009)	(0.006)	(0.008)
educ				-0.172*	-0.173*	-0.177*
				(0.076)	(0.076)	(0.076)
diabetes						-0.030
						(0.035)
handwashing						0.001
						(0.008)
Constant	6.024***	6.250***	-5.471***	-4.421**	-4.536**	-4.536**
	(0.641)	(0.494)	(1.530)	(1.434)	(1.387)	(1.396)
Observations	54	54	54	54	54	54
R2	0.050	0.152	0.550	0.707	0.706	0.711
Adjusted R2	0.013	0.101	0.513	0.669	0.676	0.667
Residual Std. Error	1.682	1.605	1.181	0.974	0.964	0.977
F Statistic	1.347	2.983*	14.966***	18.884***	23.093***	16.162***

- A country with 1USD higher HEPC, would have 0.1% more COVID deaths on average.
- A country with 1 person per sq mile more, would have 0.3% lesser COVID deaths on average.
- If a country's population increases by 1% , it would have 1.657% more COVID deaths on average.
- A country with a 1 point higher AIDI score, would have 4.3% more COVID deaths on average.
- If a country's mandatory schooling is increased by 1 year , it would have 17.3% lesser COVID deaths on average.
- All estimates are statistically significant and the magnitude of the effects is economically significant.
- HEPC, population density, log of population, AIDI, and education explain 67.6% of the variation in log deaths.

V. Regression Analysis: Alternative

COVID in Africa - Alternative Regressions

	Dependent variable:					
	log(deaths)					
	(1)	(2)	(3)	(4)	(5)	(6)
HEPC	0.001*	0.0005*	0.0004*	0.001	0.0004	0.001
	(0.0003)	(0.0002)	(0.0002)	(0.0005)	(0.001)	(0.0003)
popDensity	-0.002*	-0.003*	-0.003*	-0.002*	-0.002*	-0.003*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
logPop	1.627***	1.796***	1.728***	1.699***	1.628***	1.614***
	(0.193)	(0.222)	(0.200)	(0.184)	(0.178)	(0.176)
AIDI	0.041***	0.038***	0.038***	0.039***	0.044***	0.041***
	(0.006)	(0.007)	(0.006)	(0.007)	(0.008)	(0.006)
educ	-0.153	-0.139	-0.112	0.374	3.554	-0.161*
	(0.079)	(0.076)	(0.074)	(0.549)	(2.453)	(0.078)
dict	0.358	0.385	0.402			
	(0.285)	(0.277)	(0.271)			
polStab		0.345	0.347			
		(0.177)	(0.182)			
FDI			-0.045			
			(0.031)			
I(educ2)				-0.032	-0.386	
				(0.033)	(0.262)	
I(educ3)					0.013	
					(0.009)	
Constant	-4.679**	-5.628***	-5.239***	-7.089**	-15.713*	-4.291**
	(1.439)	(1.567)	(1.449)	(2.625)	(7.183)	(1.341)
Observations	54	54	54	54	54	53
R2	0.716	0.734	0.744	0.711	0.718	0.661
Adjusted R2	0.679	0.694	0.699	0.674	0.676	0.625
Residual Std. Error	0.959	0.937	0.929	0.966	0.964	0.968
F Statistic	19.707***	18.166***	16.386***	19.299***	16.765***	18.305***

- In the alt regression, we included dictatorship (dummy), political stability and FDI which we thought may contribute to COVID-19 deaths.
- Adjusted R2 increased, but these variables are <u>not</u> statistically significant and also decreased the estimator's coefficient.
- The baseline shows significant effect to R2 when we add education, so we added square and cube variations of this estimator, which indicated a rise in adjusted R2 but not much change in the coefficients.

V. Regression Analysis: Hypothesis Tests

```
Linear hypothesis test

Hypothesis:
I(educ^3) = 0

Model 1: restricted model
Model 2: log(deaths) ~ HEPC + popDensity + logPop + AIDI + educ + I(educ^2) +
    I(educ^3)

Note: Coefficient covariance matrix supplied.

  Res.Df Df      F Pr(>F)
1     47
2     46  1 1.9909  0.165
```

```
Linear hypothesis test

Hypothesis:
I(educ^2) = 0
I(educ^3) = 0

Model 1: restricted model
Model 2: log(deaths) ~ HEPC + popDensity + logPop + AIDI + educ + I(educ^2) +
    I(educ^3)

Note: Coefficient covariance matrix supplied.

  Res.Df Df      F Pr(>F)
1     48
2     46  2 1.1522 0.3249
```

Based on these F-tests, we can conclude that the linear specification is preferred to the quadratic and cubic ones.

VI. Data Limitations

- Limitations on recordkeeping and availability of data in developing countries, especially ones with large rural populations.
- South Africa is an outlier - it has more than triple the amount of deaths than the country with the second highest COVID-19 deaths.
- Our variables don’t cover the same period of time - COVID-19 deaths is the total # of deaths since the beginning of the pandemic until current day, while most of our variables cover the entirety of 2019 (or most recent data available).
- 8 countries did not have data for % of population that has handwashing facilities available to them.

VII. Conclusion

- Healthcare expenditure, population, Africa infrastructure development index have positive effect on COVID deaths.
- Population density and education have negative causal effect on COVID deaths in Africa.
- A country's percentage of population that's over 65 years of age or has diabetes or has access to handwashing facilities does not seem to have a causal effect on COVID deaths in Africa, which is contradictory to what has been indicated by general science, which might be due to under counting of COVID-19 deaths by country.