Exploring national COVID-19 variability across sub-Saharan Africa

Fikresus Amahazion

1 College of Arts and Social Sciences, Adi Keih, Eritrea

Keywords: pandemic, income inequality, gini index, health, covid-19, coronavirus, age, africa

https://doi.org/10.29392/001c.24941

Background

In early March 2020, coronavirus disease (COVID-19), an infectious disease caused by a novel coronavirus, was declared a pandemic by the World Health Organization. Since its emergence and global spread, the pandemic has been one of the greatest global crises in modern human history. Notably, in Sub-Saharan Africa (SSA), COVID-19-related burden and outcomes have been generally lower than many other parts of the world and substantially better than were initially feared. At the same time, there has been great heterogeneity in COVID-19 burden and outcomes between countries in the region, with some reporting particularly high incidence and death figures compared to others. What accounts for the significant cross-country variability apparent in SSA and why have some countries performed better than others? The present study investigates country-specific factors that may help to explain differences in COVID-19 outcomes across 48 countries in SSA.

Methods

A novel cross-sectional dataset, comprising a wide array of socio-demographic, political, economic, and health-related variables, is constructed through gathering data from publicly available sources. Descriptive statistics, correlation analyses, and multiple regression analyses are performed to reveal important country-level factors associated with COVID-19 deaths in SSA.

Results

Findings from statistical analyses show that in SSA COVID-19 deaths per million is positively associated with income inequality and median age, and negatively associated with population density. In contrast, a number of other variables, including gross national income (GNI) per capita, global connectivity, diphtheria, tetanus and pertussis (DTP) immunization coverage, the proportion of seats in parliament held by women, and political system or regime type, are not statistically significant.

Conclusions

Although findings from recent studies conducted in various settings around the world indicate that a range of socio-economic, demographic, political, and health-related factors may be linked with COVID-19 burden, the present investigation finds that COVID-19 deaths in SSA are associated with population density, median age, and income inequality.

Coronavirus disease (COVID-19) is an infectious disease caused by a novel coronavirus. In early March 2020, COVID-19 was declared a pandemic by the World Health Organization.1 Since its emergence and global spread, the pandemic has been one of the greatest crises in decades, with over one hundred million confirmed cases, several million deaths, and an array of socio-economic, political, and other impacts. Although COVID-19 has been a truly global problem, affecting all regions, the COVID-19 burden in Sub-Saharan Africa (SSA) has been substantially lower than was predicted. Countries in SSA reported first confirmed cases and experienced a surge of the virus much later than other parts of the world, while also consistently reporting a comparatively low number of cases and deaths.2,3 Additionally, the case-fatality ratio for many countries in SSA has generally been lower than many other parts of the world, suggesting that disease outcomes have been less severe among populations in the region.4,5

At the time, while SSA has had a generally lower overall COVID-19 burden than many other parts of the world, there has been great heterogeneity in burden and outcomes between individual countries within the region. For instance, some countries, such as South Africa and Ethiopia, have reported considerably higher prevalence and mortality com-
pared to others. What accounts for the significant cross-country variability apparent in SSA and why have some countries performed better than others? The present study examines factors potentially associated with COVID-19 burden in the region. Through conducting a multiple regression analysis on a unique cross-sectional dataset comprising 48 countries from SSA, important country-level factors associated with COVID-19 deaths are investigated.

The present study is significant for several reasons. It strengthens and deepens understanding of prominent factors associated with COVID-19 burden, thus complementing the existing knowledge base and potentially informing policies or approaches to future pandemics. As well, whereas many previous studies have focused on biological and medical factors related to COVID-19, this study extends consideration to socio-economic, political, and other country-level factors which can also play a role in elucidating differences in COVID-19 outcomes across countries. Moreover, it investigates a considerably large number of countries and factors potentially associated with COVID-19, therefore offering a broader, wide-ranging, and more comprehensive investigation than many previous studies, which have been limited to examining single or only a few factors or countries. Last, the present study expands awareness and offers useful insights about COVID-19 in SSA, a massive region comprising 48 countries, which has received comparatively limited coverage and a paucity of research attention.

The paper is structured as follows. The methods are presented in the next section, featuring an outline of the design and description of the data. Subsequently, the results are reported, followed by the discussion. The final section presents conclusions and notes possible limitations.

METHODS

DATA

The present study is a cross-sectional analysis of 48 countries from SSA, as classified by the World Bank’s regional designations. A unique dataset is constructed utilizing data that was extracted from various public open access databases. The variables selected for analysis cover sociodemographic, political, economic, and health-related dimensions of countries and are included on the basis of empirical findings reported in recent studies, the growing and evolving literature, relevant theoretical frameworks, and general data availability.

VARIABLES

COVID-19 burden: consistent with a large number of empirical studies, cumulative total deaths per million of the population is the outcome variable. The data is from the Our World in Data open access database and accurate up to 7 May 2021.

Gross national income per capita: this variable presents a country’s gross national income (GNI), divided by its total population. Data are presented in current US dollars and taken from the World Bank’s World Development Indicators (WDI) open access database, which provides comprehensive cross-country comparable data on development.

Global connectivity: the KOF Globalisation Index (KOFGI), which is the most widely used globalization index in the academic literature, is used as a proxy for global connectivity. A composite index based on 43 individual variables, the KOFGI measures globalization for every country in the world along the economic, social, and political dimension. Higher KOFGI values reflect greater degrees of globalization.

Health system: data for traditional measures of national health system capacity (e.g., physicians per 1,000 population, intensive care unit capacity, or hospital beds per 1,000 population) is unavailable for many countries. Accordingly, national diphtheria, tetanus, and pertussis (DTP) immunization coverage is utilized. DTP coverage is regarded as a useful, standard measure of the strength and efficiency of national health systems, mainly because delivery requires three contacts with the health system at appropriate times and also because coverage is usually part of routine national vaccination programs rather than campaigns. Defined as the percentage of children ages 12-23 months who received vaccinations before 12 months or at any time before the survey, this variable is available from the WDI database.

Population density: drawn from the WDI dataset, population density is measured as the total population of a country divided by its total land area in square kilometers.

Median age of population: this variable, gathered from the World Population Prospects public database, presents the median age of the population of a country, which divides the population in two parts of equal size, so that there are as many persons with ages above the median as there are with ages below the median. It is expressed in years.

Polity: the data for this collected from extremely popular and extensively used Polity5 Project database, which codes the political regime and authority characteristics of states in the world system for purposes of comparative,
The results of the regression analyses are presented. The different variables are included in several stages, according to socio-demographic, political, eco-

### Table 1. Summary statistics of variables for sub-Saharan African countries

<table>
<thead>
<tr>
<th>Variable</th>
<th>Observations</th>
<th>Mean</th>
<th>Maximum</th>
<th>Minimum</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deaths per million</td>
<td>48</td>
<td>93.39</td>
<td>922.07</td>
<td>0.35</td>
<td>168.39</td>
</tr>
<tr>
<td>Population density</td>
<td>48</td>
<td>108.84</td>
<td>623.30</td>
<td>2.97</td>
<td>137.19</td>
</tr>
<tr>
<td>Median age</td>
<td>48</td>
<td>20.05</td>
<td>37.50</td>
<td>15.20</td>
<td>4.19</td>
</tr>
<tr>
<td>GNI per capita</td>
<td>45</td>
<td>2443.56</td>
<td>16900.00</td>
<td>280.00</td>
<td>3314.23</td>
</tr>
<tr>
<td>Inequality</td>
<td>45</td>
<td>43.81</td>
<td>63.00</td>
<td>32.60</td>
<td>7.81</td>
</tr>
<tr>
<td>Globalization</td>
<td>47</td>
<td>50.15</td>
<td>72.00</td>
<td>30.00</td>
<td>8.44</td>
</tr>
<tr>
<td>DTP coverage</td>
<td>48</td>
<td>80.31</td>
<td>99.00</td>
<td>42.00</td>
<td>16.34</td>
</tr>
<tr>
<td>Polity</td>
<td>48</td>
<td>2.81</td>
<td>10.00</td>
<td>9.0</td>
<td>4.91</td>
</tr>
<tr>
<td>Women in parliament</td>
<td>47</td>
<td>21.70</td>
<td>61.25</td>
<td>3.38</td>
<td>12.47</td>
</tr>
</tbody>
</table>

DTP – diphtheria, tetanus and pertussis, GNI – gross national income

For preliminary exploration of the relationships between variables, summary statistics and correlations are examined. Correlation analysis explores the association between variables, allowing for better understanding of the magnitude and direction of relationships. Correlation coefficients are measured on a standard scale that ranges between -1 and +1, with positive and negative values indicating the direction of relationship. Coefficients that are closer to 0 reflect a very weak or small association. As well, regression analysis, one of the most widely used quantitative analysis. For each country, a "Polity Score" is determined, ranging from +10 (strongly democratic) to -10 (strongly autocratic). Scores are based on an evaluation of democratic and autocratic characteristics and elements of regimes, including elections, the nature of political participation, and the extent of checks on executive authority. Women's leadership and parliamentary representation: this variable is the number of seats held by women members in single or lower chambers of national parliaments, expressed as a percentage of all occupied seats. Retrieved from the WDI dataset, it is derived by dividing the total number of seats occupied by women by the total number of seats in parliament. Notably, regression analysis has been used in a considerable amount of work conducted on COVID-19. All analyses were performed utilizing the Statistical Package for the Social Sciences (23rd ed.) statistical software program.

### RESULTS

#### SUMMARY STATISTICS

The summary statistics are shown within Table 1. The mean cumulative COVID-19 deaths per million is 93.59, mean GNI per capita is 2443.56, and mean global connectivity value is 50.15. For inequality, the mean is 45.81, while the means for DTP coverage and polity score are 80.31 and 2.81, respectively. The mean median age is 20.05, while population density has a mean of about 108.84. Last, the mean percentage of parliamentary seats held by women is 21.70.

#### CORRELATION COEFFICIENTS

Pearson's correlation coefficients, which help to reveal possible associations between variables and COVID-19 burden, are reported in Table 2. The results illustrate that there is a moderate, statistically significant positive correlation between COVID-19 deaths per million and GNI per capita ($r = .398, P = .007$), global connectivity ($r = .378, P = .009$), inequality ($r = .551, P = .000$), and median age ($r = .471, P = .001$). This indicates that higher GNI per capita and greater inequality are associated with more COVID-19 deaths per million. Furthermore, as the value of the median age and globalization of a country increases, so does the value of cumulative COVID-19 deaths per million.

#### REGRESSION ANALYSIS

In Table 3, the results of the regression analyses are presented. The different variables are included in several stages, according to socio-demographic, political, eco-
nomic, and health-related dimensions, in order to better examine their possible association with COVID-19 deaths per million.

Panel I explores population density and median age. Results reveal that population density and median age together explain 29.4% of the variance ($R^2=0.294$) and that Panel I is a significant predictor of deaths per million, $F(2, 45)=9.371, P<0.001$. Both population density ($B=-0.357, P<0.001$) and median age ($B=23.444, P<0.001$) contribute significantly.

In Panel II, population density and median age are retained, while GNI per capita and the measure of inequality are introduced. Panel II accounts for approximately 50.4% of the variability in deaths per million ($R^2=0.504$) and is significantly useful in explaining deaths per million $F(4, 39)=9.896, P<0.001$. The estimates show that median age ($B=29.4, P=0.006$) and inequality ($B=10.173, P<0.001$) are positively associated with deaths per million. That is, as median age increases by one unit, cumulative COVID-19 deaths per million increase by 29.4 units, while inequality increases by one unit, COVID-19 deaths per million increase by 10.173 units. However, population density is negatively associated with deaths per million ($B=-.285, P=0.078$) at the 10% level, while GNI per capita fails to demonstrate a statistically significant relationship with COVID-19 deaths per million.

Panel III builds on the previous panels by adding DTP coverage and global connectivity. About 51.2% of the variability in COVID-19 deaths per million can be accounted for with Panel III ($R^2=0.512$), which is also significantly useful in explaining deaths per million $F(6, 37)=6.466, P=0.001$. Once again, the estimates show that median age ($B=26.572, P=0.020$) and inequality ($B=10.048, P<0.001$) are positively associated with COVID-19 deaths, while population density ($B=-0.303, P=0.085$) remains negatively associated at the 10% level. However, the various other variables included do not make a statistically significant contribution.

In Panel IV all of the variables are included. Panel IV explains approximately 54.2% of the variation in COVID-19 deaths per million ($R^2=0.542$) and is significantly useful in explaining COVID-19 deaths per million $F(8, 34)=5.033, P=0.000$. As with other regressions, median age ($B=51.991, P=0.011$) and inequality ($B=9.442, P=0.002$) demonstrate a positive association with COVID-19 deaths per million, while population density also contributes significantly, being negatively associated with deaths per million at the 5% level ($B=-0.384, P=0.049$). Thus, as median age and inequality increase, so do COVID-19 deaths per million. On the other hand, as population density increases, cumulative COVID-19 deaths per million decrease. The various other variables included do not make a statistically significant contribution.

**DISCUSSION**

The findings in the present study are noteworthy and provide several key insights. For instance, past scholarship has demonstrated an association between inequality and various health outcomes, while recent studies from various settings show an association between inequality and COVID-19 burden. Consistent with this body of work, the present study finds a clear positive association between inequality and COVID-19 deaths.

There are several different possible mechanisms that may help explain this finding. Societies with greater levels of inequality may have less social cohesion or lower public trust levels, which can undermine compliance with government and public health guidelines and lifestyle factors, such as diabetes, obesity, and lack of exercise, which can leave individuals or communities vulnerable to infections, severe cases, and deaths. Inequality may also increase risk of infection, as the most disadvantaged individuals cannot work from home and must remain in high-risk employment, while inequality and exclusion are associated with differences in health literacy and information seeking behavior, which can impact health outcomes.

Also, the findings related to population density are notable. In particular, previous work has illustrated that population density can be a major factor in the transmission of infectious diseases and that epidemics or outbreaks may be more frequent or severe when population density is high. More recently, a substantial body of work has found an association between population density and COVID-19. In stark contrast, this study finds that population density is negatively associated with COVID-19 deaths. While potentially somewhat surprising, this is in accord with recent work indicating that higher population density may be linked with fewer COVID-19 cases or deaths. Possible explanations include that areas of greater density generally have better and more concentrated healthcare infrastructure and facilities, more quality staff and specialized services, faster response times, and more intensive-care

**Table 2. Pearson correlation coefficients**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Deaths per million</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population density</td>
<td>-.063 (.669)</td>
</tr>
<tr>
<td>Median Age</td>
<td>.471*** (.001)</td>
</tr>
<tr>
<td>GNI per capita</td>
<td>.398*** (.007)</td>
</tr>
<tr>
<td>Inequality</td>
<td>.531*** (.000)</td>
</tr>
<tr>
<td>Globalization</td>
<td>.378*** (.009)</td>
</tr>
<tr>
<td>DTP coverage</td>
<td>.190 (.195)</td>
</tr>
<tr>
<td>Polity</td>
<td>.098 (.507)</td>
</tr>
<tr>
<td>Women in parliament</td>
<td>.156 (.296)</td>
</tr>
</tbody>
</table>

P-values are in parentheses.

*** = Correlation is significant at the 0.01 level.

** = Correlation is significant at the 0.05 level.

* = Correlation is significant at the 0.10 level.

DTP – diphtheria, tetanus and pertussis, GNI – gross national income.
Table 3. Regression analysis of factors associated with COVID-19 deaths in Africa

<table>
<thead>
<tr>
<th>Panel</th>
<th>Population density</th>
<th>Median age</th>
<th>GNI per capita</th>
<th>Inequality</th>
<th>Globalization</th>
<th>DTP coverage</th>
<th>Polity</th>
<th>Women in parliament</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>-.357** (.038)</td>
<td>23,444*** (.000)</td>
<td>-0.014 (.286)</td>
<td>10.173*** (.000)</td>
<td>1.625 (.666)</td>
<td>.846 (.605)</td>
<td>-7.943</td>
<td>- .330 (.853)</td>
</tr>
<tr>
<td>II</td>
<td>-.285* (.078)</td>
<td>29,400*** (.006)</td>
<td>-0.013 (.314)</td>
<td>10.048*** (.001)</td>
<td>1.625 (.666)</td>
<td>.846 (.605)</td>
<td>-7.943</td>
<td>- .330 (.853)</td>
</tr>
<tr>
<td>III</td>
<td>-.303* (.085)</td>
<td>26,572** (.020)</td>
<td>-0.019 (.190)</td>
<td>9.442*** (.002)</td>
<td>1.047 (.540)</td>
<td>.542</td>
<td>- .330 (.853)</td>
<td></td>
</tr>
<tr>
<td>IV</td>
<td>-.384** (.049)</td>
<td>31,991** (.011)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Dependent variable is COVID-19 deaths per million.
P-values are in parentheses.
*** = significant at the 0.01 level.
** = significant at the 0.05 level.
* = significant at the 0.10 level.
DTP – diphtheria, tetanus and pertussis, GNI – gross national income.

From early on in the COVID-19 pandemic, researchers have widely documented that older populations, who are more likely to have comorbidities, declining immune systems, and preexisting medical conditions, have more vulnerability to infection, dramatically greater risk of developing severe and more serious complications, and strikingly higher mortality rates compared to younger groups.37,38 Furthermore, studies have found a strong association between median age and COVID-19 burden. The present study confirms this existing body of empirical work, finding a clear association between median age and COVID-19 deaths.

Global connectivity and levels of international air traffic or travel facilitated the spread of disease during past crises, and some research suggests that they may have also played a role in the early importation of COVID-19.40,41 Still, the present study finds that global connectivity, as measured by a standard, extensively used index of globalization, the KOFGI, is not associated with COVID-19 deaths.

Generally, national wealth, income, or economic performance, and healthcare systems or capacity have consistently been shown to be major determinants of health and development outcomes.42 However, the fact that GNI per capita and DTP coverage do not demonstrate statistical significance in the present study parallels work suggesting that countries’ level of wealth, development, and income is not the main criteria or sole determinant of health outcomes or COVID-19 burden. It is quite notable that as the COVID-19 pandemic has unfolded around the world, many of the wealthiest, highly developed countries, characterized by well-equipped healthcare systems and considerable, high-quality resources, have been ill-prepared and among those most severely hit. On the other hand, some less developed, low- and middle-income countries have reported much lower COVID-19 incidence and mortality rates. Despite facing genuine obstacles and various constraints, the latter may be able to plan and execute timely, innovative, and efficient responses, coordinate efforts of various sectors and mobilize civil society, implement and enforce effective policies or guidelines, and maximize limited assets or resources.43

A broad literature exists exploring the relationship, vulnerability, and response of different political systems or regime types to disasters, emergencies, and crises.44 Since the onset of the COVID-19 pandemic, as well, the questions of whether and how different political systems may impact national COVID-19 burden have been prominent and widely debated. While the scholarship continues to grow and develop, to date, there is no strong or unambiguous consensus; there are conflicting arguments and findings apparent and both forms of government seem to have specific strengths and weaknesses.45 This study does not find an association between political system and COVID-19 deaths. The pandemic has not distinguished between democracies and autocracies in SSA. Moreover, there has been wide variance in responses within democratic and autocratic regimes.
in the region, and strictly focusing on the type of political system risks neglecting other potentially significant factors. Rather than the particular type of political regime, findings suggest that several country-specific variables, such as income inequality, population density, and age, have more bearing on COVID-19-related outcomes.

Another frequent topic of commentary, debate, and comparison during the COVID-19 pandemic has been whether countries led by women have fared better than those led by men. Plausibly, gender may be significant in different ways. For example, some work has indicated that women tend to be more loss- and risk-averse than men, therefore possibly making them less willing to accept health risks and more likely to introduce restrictive measures earlier. Additionally, more women in legislature has also been associated with an increase in public health spending, while a lack of diverse representation in leadership can exclude those who offer unique perspectives or expertise. Johnson and Williams also suggest that women leaders may be able to draw on their traditional motherly role - for example, as the member of the household who traditionally cares for the sick - to display forms of feminine protectionism.

Notably, in a study of COVID-19 deaths in the United States, it was found that states with women governors had fewer COVID-19 deaths compared to states with men governors. Similarly, in a study of differences in the policy responses of male and female leaders from 194 countries, researchers found that COVID-19 outcomes, especially deaths, are better in countries led by women, who reacted more quickly and decisively in the face of potential fatalities.

Nevertheless, the present study finds that the proportion of seats held by women in national parliaments is not statistically significant, thus generally aligning with other empirical studies that have found no statistical support for differences in reported fatalities between women-led and men-led countries notion. It might be the case that rather than leadership gender, what matters most for COVID-19-related outcomes are the specific features and characteristics of the countries that they oversee and administer.

CONCLUSION

The COVID-19 pandemic has been one of the greatest global crises in modern human history. Interestingly, in SSA, COVID-19-related outcomes have been generally better than were initially feared or expected. At the same time, there has been great heterogeneity in COVID-19 outcomes between countries in the region, with some reporting particularly high burdens compared to others. Using a novel dataset comprising an array of socio-demographic, political, economic, and health-related variables gathered from publicly available sources, the present study investigates country-specific factors that may help to explain differences in COVID-19 outcomes across 48 countries in SSA. Findings from statistical analyses demonstrate that COVID-19 deaths per million is positively associated with inequality and median age, and negatively associated with population density. On the other hand, a number of other variables, including GNI per capita, global connectivity, DTP immunization coverage, the proportion of seats in parliament held by women, and political system or regime type, are not statistically significant.

A number of important limitations in the present study can be noted. For one, the findings are dependent on primary data sources, some of which may be of uncertain quality. In particular, it is possible that COVID-19 data for some countries in SSA are not wholly accurate, since infections and deaths may be underreported due to limited testing capacity and challenges in the attribution of the cause of death (e.g., weak or nonexistent national death registration systems). Thus, findings and conclusions, while informative and enlightening, must also be considered with care and caution.

Second, despite many countries around the world making considerable progress in the rollout of rapidly developed vaccines, SSA remains considerably far behind. The pandemic continues to evolve in the region, new and highly transmissible variants are circulating, and future dynamics are still yet unclear. As a result, some of the findings and conclusions presented here may possibly change. Additionally, while the present study examined many country-specific variables, there are still several others that may be included in future analyses, including the role of civil society, experience with past epidemics or crises, influence of media or information, and temperature or climate. Last, the present study is cross-sectional, thus findings may be different for other timeframes or when utilizing time series, longitudinal data.

Notwithstanding these limitations, however, the present study contributes to our general understanding of important factors associated with COVID-19 burden and helps to expand awareness of the regional situation in SSA, which has received comparatively limited coverage and a paucity of research attention. Moreover, rather than representing a final, definitive judgment or account of the situation in the region, the findings in the present study offer useful insights and indicate possible avenues for further debate, research, and investigation.

ETHICS

No protocol approval was required because no research involving human participants was conducted.

FUNDING

There are no sources of funding to declare.

COMPETING INTERESTS

The authors completed the Unified Competing Interest form at www.icmje.org/coi_disclosure.pdf (available upon request from the corresponding author) and declare no conflicts of interest.

CORRESPONDENCE TO:

Dr. Fikresus Amahazion
College of Arts and Social Sciences
REFERENCES


