Who Did Covid-19 Hurt the Most in Sub-Saharan Africa?

Feraud Tchuisseu Seuyong
Ifeanyi Edochie
David Newhouse
Ani Rudra Silwal
Abstract

How did the economic crisis caused by the Covid-19 pandemic impact poor households in Sub-Saharan Africa? This paper tackles this question by combining 73 High-Frequency Phone Surveys collected by national governments in 14 countries with older nationally representative surveys containing information on household consumption. In particular, it examines how outcomes differed according to predicted per capita consumption quintiles in the first wave of the survey, and in subsequent waves by households' predicted per capita consumption. The initial shock affected households throughout the predicted welfare distribution. Households in the bottom 40 percent responded by sharply increasing farming activities between May and July of 2020 and gradually increasing ownership of non-farm enterprises starting in August. This coincided with an improvement in welfare, as measured by a decline in food insecurity and distressed asset sales among these households during the second half of 2020. With respect to education, children in the bottom quintile were 15 percentage points less likely to engage in learning activities than those in the top quintile in the immediate aftermath of the crisis, and the engagement gap between the bottom 40 and top 60 widened in the summer before narrowing in the fall due to large declines in engagement among the top 60. Poorer households were slightly more likely to report receiving public assistance immediately following the shock, and this difference changed little over the course of 2020. The results highlight the widespread impacts of the crisis both on welfare and children's educational engagement, the importance of agriculture and household non-farm enterprises as safety nets for the poor, and the substantial recovery made by the poorest households in the year following the crisis.

This paper is a product of the Poverty and Equity Global Practice and the Development Data Group, Development Economics. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://www.worldbank.org/prwp. The authors may be contacted at dnewhouse@worldbank.org. A verified reproducibility package for this paper is available at http://reproducibility.worldbank.org, click here for direct access.
Who did Covid-19 hurt the most in Sub-Saharan Africa?

Feraud Tchuisseu Seuyong
Ifeanyi Edochie
David Newhouse
Ani Rudra Silwal

JEL codes: O12, I31
Keywords: Covid-19, Sub-Saharan Africa, Distributional impacts
Introduction

How did the economic impact of the crisis caused by Covid-19 impacted poor households in Sub-Saharan Africa? While statistics on infections and deaths suggest that Sub-Saharan Africa (SSA) may have escaped the worst of Covid-19, the economic cost was major. Poverty in the region was projected to increase by 2 percentage points or more in 2020 (Montes et al. 2020 and Lakner et al. 2021). While the full extent of the economic damage is still unknown, data from phone surveys collected since the pandemic provide an indication of the adverse impact on livelihoods in the region. However, little is known about how impacts differed for poor and wealthy households within countries. This note aims to help fill this gap in order to better understand the distributional impacts of the crisis and inform recovery efforts.

We combine data collected through 73 phone surveys in SSA since the early stages of the global pandemic with richer multi-topic household surveys collected prior to the pandemic to examine how the impact of Covid-19 differed for poorer and wealthier households, as measured by predicted household per capita consumption. There are six main findings:

1. The initial impact of Covid-19 on the prevalence of income declines and the ability to work as usual was widespread throughout the predicted consumption distribution.
2. During the summer of 2020, there were large (20 pp) increases in the share of households in the bottom 40 percent that engaged in farming activities, and a more gradual increase 10 pp increase in the share of these households that owned or partially owned a non-farm enterprise.
3. During the latter part of 2020 and early 2021, food insecurity and the incidence of distressed asset sales declined faster for the bottom 40 than the top 60 percent. Conversely, the prevalence of income declines and the inability to work as usual declined faster for the top 60 percent, who were better able to cope with the medium-term impacts of the crisis.
4. Children in the bottom 40 percent were less likely to engage in educational activities in the immediate aftermath of the crisis. This disadvantaged narrowed greatly over the course of December 2020, as the share of children engaging in learning or educational activities fell about 10 percentage points for the bottom 40 and about 17 pp for the top 60 both between June and December 2020.
5. Public government assistance slightly favored poorer households in the spring and summer of 2020, a pattern that changed little as the crisis evolved.

Overall, the results indicate that the initial impacts of the Covid-19 were widespread throughout the predicted welfare distribution, and that the agricultural sector and establishment of non-farm enterprises acted as an important safety net. The movement back to agriculture may have helped poorer households reduce food insecurity and distressed asset sales, but may have delayed their ability to stem declines in income. Households faced serious challenges maintaining their children’s engagement with educational activities during the summer, and households in the bottom 40 were slightly more likely than those in the top 60 to benefit from public assistance throughout 2020.

---

Data and methodology

High Frequency Phone Surveys and Global Monitoring Database

This section describes the data sources and the methodology used to examine the impact of Covid-19 across the welfare distribution in 14 countries in Sub-Saharan Africa. We rely on two different databases for our analysis: High Frequency Phone Surveys (HFPS) implemented since the pandemic began and the Global Monitoring Database (GMD), a database of household surveys implemented prior to the pandemic, compiled and managed by the World Bank. We rely on survey-to-survey imputation, which entails developing a model of household consumption in the GMD for each country, and using the model to impute the welfare status of households in each wave of that country’s HFPS.

Since Spring 2020, the World Bank has supported national statistical offices in developing countries to design and implement phone surveys to examine various aspects of Covid-19 on lives and livelihoods. The rapid pace at which the Covid-19 pandemic has affected countries has made access to timely knowledge on the impacts of Covid-19 critical to effectively design, target and evaluate policy interventions. However, implementing face-to-face interviews was severely restricted in the context of widespread government-imposed social distancing practices. Phone surveys filled this gap by eliciting information from individuals and households rapidly and at low cost. In all 14 of the countries we consider, phone surveys were implemented as follow-up surveys from previous nationally representative surveys. These previous surveys were used to reweight the phone survey to make it more representative of the national population of households. However, the phone surveys disproportionately interview household heads, meaning that individual level characteristics such as the ability to work at usual were not representative of the underlying population of workers. Nonetheless, phone surveys can be helpful in tracking the responses to and impacts of the pandemic, particularly if they are combined with traditional household surveys. Table 1 lists the countries and waves of HFPS that were used for this analysis.

Table 1: Sample size of HFPS waves and the corresponding multi-topic survey from GMD

<table>
<thead>
<tr>
<th>Country</th>
<th>Wave 1</th>
<th>Wave 2</th>
<th>Wave 3</th>
<th>Wave 4</th>
<th>Wave 5</th>
<th>Wave 6</th>
<th>Wave 7</th>
<th>Wave 8</th>
<th>Wave 9</th>
<th>Wave 10</th>
<th>Wave 11</th>
<th>Survey</th>
</tr>
</thead>
<tbody>
<tr>
<td>Burkina Faso</td>
<td>1,968</td>
<td>2,210</td>
<td>2,125</td>
<td>2,104</td>
<td>2,071</td>
<td>2,032</td>
<td>2,011</td>
<td>1,998</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ethiopia</td>
<td>3,249</td>
<td>3,107</td>
<td>3,058</td>
<td>2,878</td>
<td>2,770</td>
<td>2,753</td>
<td>2,536</td>
<td>2,222</td>
<td>2,077</td>
<td>2,178</td>
<td>1,982</td>
<td>2018 EHCVM</td>
</tr>
<tr>
<td>Gabon</td>
<td>1,656</td>
<td>1,507</td>
<td>1,402</td>
<td>1,402</td>
<td>1,378</td>
<td>1,308</td>
<td>1,297</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gambia</td>
<td>1,436</td>
<td>1,385</td>
<td>1,335</td>
<td>1,332</td>
<td>1,332</td>
<td>1,297</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mali</td>
<td>1,766</td>
<td>1,935</td>
<td>1,901</td>
<td>1,901</td>
<td>1,901</td>
<td>1,884</td>
<td>1,760</td>
<td>1,694</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mozambique</td>
<td>1,097</td>
<td>978</td>
<td>960</td>
<td>841</td>
<td>901</td>
<td>901</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Malawi</td>
<td>1,729</td>
<td>1,646</td>
<td>1,624</td>
<td>1,616</td>
<td>1,589</td>
<td>1,591</td>
<td>3,112</td>
<td>3,098</td>
<td>3,090</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Niger</td>
<td>1,264</td>
<td>1,264</td>
<td>1,264</td>
<td>1,264</td>
<td>1,264</td>
<td>1,264</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nigeria</td>
<td>1,950</td>
<td>1,820</td>
<td>1,790</td>
<td>1,789</td>
<td>1,773</td>
<td>1,762</td>
<td>1,726</td>
<td>1,723</td>
<td>1,706</td>
<td>1,699</td>
<td>1,680</td>
<td>2018 EHCVM</td>
</tr>
<tr>
<td>Senegal</td>
<td>1,220</td>
<td>1,220</td>
<td>1,220</td>
<td>1,220</td>
<td>1,220</td>
<td>1,220</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A secondary source of data is the Global Monitoring Database (GMD), which is a collection of globally harmonized household survey data compiled by the World Bank. The GMD includes most recent household surveys used by national statistical agencies around the world, and subsequently, by the World Bank, to compute the official poverty statistics. In addition to the welfare aggregate, surveys in GMD include variables on demographics, assets, labor, education, and other topics that have been harmonized to a common set of specifications, making the GMD well-suited for cross-country analysis.

This analysis covers 14 countries in SSA (Table 1). Although more countries are included in both the GMD and HFPS, we were limited by three factors. First, the analysis could only be conducted in countries where the HFPS was conducted and which also has a survey available in GMD that was conducted prior to the pandemic. Second, we limited our analysis only to surveys in GMD that were conducted in recent years, using a cutoff of 2014. Third, we limited our analysis to countries that both had a minimum set of variables that could be used to model household per capita consumption.

Survey-to-survey imputation

Survey-to-survey imputation (S2S) predicts household per capita consumption in a household survey that collects information only on selected covariates of consumption by using information from another survey that collected both the per capita consumption and its covariates. This literature builds largely on the poverty mapping literature (Elbers, Lanjouw, and Lanjouw, 2003) and has been widely implemented (Dang 2020, Newhouse et al. 2014, Stifel and Christiaensen 2007). This literature is in contrast to the multiple imputation literature inspired by Rubin (1987), which largely focuses on filling missing data within a survey rather than using information from one survey to make predictions in another one. There are generally two types of S2S studies in the literature: one that conducts the imputation over time and another that conducts the imputation within the same period but across surveys. The first type includes studies such as those by Newhouse and Vyas (2022) and Stifel and Christiaensen (2007). Studies that impute across surveys within the same time frame include Beltram et al. (2020), Dang (2019), Lucchetti et al. (2018), Dang et al. (2018), and Cuesta and Ibarra (2017). Dang (2020) and Dang et al. (2019) provide guidance on the appropriate methodology and caveats to using S2S.

Implementation of S2S in this study first involved identifying a set of variables that are available both in HFPS and GMD. This yielded the following variables: household size, urban/rural location, subnational administrative area, share of children, and share of elderly residents in the household. This list of variables was selected by carefully comparing the data dictionaries of HFPS and GMD. Both the HFPS and the GMD are both designed to be nationally representative, except for Mozambique, where the HFPS only covers urban areas. A major difference between the HFPS and the GMD is that while the GMD sample covers both households that own phones and those that do not, the HFPS (by design) only covers households that own a mobile phone. Since households that own mobile phones are likely to be a non-random sample of the population, we select a comparable sub-sample of mobile phone owners in the GMD for the purposes of S2S. Figure 1 compares these samples for the 14 countries in the sample. Appendix Figure 1 presents these differences for individual countries in our analysis sample.
As mentioned above, the HFPS does not randomly sample respondents within the household, and we therefore mainly considered household-level characteristics for imputation. The only individual level characteristic considered is the ability to work as usual. In particular, respondents in the HFPS are more likely to be household heads, older, male, and urban. This implies that modeling consumption using individual variables may yield a consumption distribution (in HFPS) that is significantly different than the national distribution in GMD. Therefore, we use survey-to-survey models specified at the household level.

Figure 1: Comparison of variables between GMD and HFPS Wave 1

![Figure 1: Comparison of variables between GMD and HFPS Wave 1](image)

Note: Table shows the mean or median of the differences across the 14 countries in the sample between the full GMD, the GMD with phone owners only, and the HFPS. Appendix Figure 2 presents numbers for all countries.

The next step in the S2S procedure involved building a model of household per capita consumption as a function of the variables (and their interaction terms) that are common between the GMD and the HFPS. The Least Absolute Shrinkage and Selection Operator (LASSO) technique was implemented to select a subset of all the candidate variables, to avoid overfitting the model to the specific GMD sample. The coefficients of the (post-)lasso model were then used to predict per capita household consumption of all households in the HFPS across all waves. The S2S methodology implemented with the adjusted R-squared estimated for each country are described in more details in the appendix. The section also compares the observed and predicted distributions of per capita household in the GMD with the predicted consumption in the HFPS. We then generate quintiles of the predicted per capita household consumption in the HFPS to tabulate various outcomes across the quantiles. Once the quintiles are generated, we perform two sets of analysis: The first looks at differences across quintiles in the initial wave of the survey, while the second looks at the average differences between the bottom 40 and top 60 by month. The next section describes the main findings of this analysis.

Tracking the evolution of the crisis

---

3 The only individual level characteristic considered is the ability to work as usual.
4 When implementing LASSO, we selected the penalty parameter lambda that minimize the Bayesian Information Criterion (Zhang, Li, and Tsi 2010). We ensured that if an interaction term was selected by LASSO, the appropriate level terms were also included in the final model, following Chen, Li, and Wang (2020). A Stata ado file that implements this is available from the authors upon request.
5 Observations were weighted using household survey weights when constructing the quintiles.
With a measure of household predicted welfare in hand, it is straightforward to report averages of outcomes from the first survey wave by predicted welfare quintile, to get a sense of how the initial impacts of the crisis varied over the predicted per capita consumption distribution. Analyzing how these differences evolved over time, however, is more complicated because countries enter and exit the sample, forming an unbalanced sample. To analyze the evolution of the impact of Covid-19, we therefore pooled data from 73 phone surveys from the 14 countries with imputed welfare. The goal of the analysis is to investigate how trends in income loss, employment, food insecurity, and social protection evolved for households in the bottom two quintiles as opposed to the top three. To examine trends, we estimate a logit model using the following model specified at the household level:

\[ y_{htc} = \beta_1 f(t) + \beta_2 f(t) \ast B40_{htc} + \theta_c + \epsilon_{htc} \]

\( y_{htc} \) is the outcome of interest observed in household \( h \), region (i.e. state/province) \( r \), country \( c \), and month \( t \). In this case it is in this case is one of eight indicators: “Household received decreased total income”, “Respondent could work as usual”, “Household with children engaged in any learning/education activities since school closures”, “Households with adult member skipped a meal in the last month due to lack of money or other resources”, “Household engaged in farming activities since the beginning of 2020”, “Households owning a non-farm enterprise”, “Households receiving government assistance”, and “Households that sold assets to cope with the pandemic”. \( B40_{htc} \) is a dummy variable taking on one if household \( h \) is in the bottom 40 percent. \( \theta_c \) is a country fixed effect, and \( U_{htc} \) is a dummy variable indicating that household \( h \) resides in an urban area. The term \( \beta_1 f(t) \) is the cubic polynomial in the month of the survey, measured since March 2020. Thus, the \( \beta_1 f(t) \) term reflects the average trend across months for the top three quintiles and the \( \beta_1 f(t) + \beta_2 f(t) \ast B40_{htc} \) term represents the average trend for households in the bottom 40 percent. Finally, \( \epsilon_{htc} \) is the error term, clustered at the level of the survey wave. After estimating the model, we plot the average predictive margins for the bottom 40 and top 60 percent by month, evaluated at the means of the other covariates and treating all dummy variables as balanced.

**Results**

This section contains the main findings of this analysis based on a S2S imputation of per capita household computation in HFPS based on a model of consumption in GMD. We examine the impact of Covid-19 on changes in household income and labor, food security, children’s education, and coping strategies implemented by households across the consumption distribution.

**Household Income and labor**

The initial impact of Covid-19 on the prevalence of reported total income declines was widespread, with minor differences across the distribution. Figure 2a plots the share of households that experienced a decrease in total income immediately after Covid-19 struck. Across the 14 countries analyzed in this sample, the average country reported that more than 76% of households experienced a decrease in total income. This share is about 5 percentage points higher for the poorest quintile (Q1). Appendix Figure 3 show a similar pattern for farm income. Declines in wage income, in the initial wave of the HFPS, were most prevalent in the third and fourth quintile. Figure 2c describes how the prevalence of total income declines has evolved since the beginning of the pandemic, after controlling for country and urban location.
Through the first six months, households in the top 60 percent of the welfare distribution were equally likely to report total income declines since the previous survey wave than the bottom 40 percent. Around November 2020, the bottom 40 percent became 5 to 10 percentage points more likely to report income declines, with a statistically significant difference in February 2021.

Figure 2: The impact of Covid-19 on income and labor across the welfare distribution

Figure 2a: Households that experienced a decrease in total income (wave 1)

Figure 2b: Share of Individuals that stopped working (wave 1)

Figure 2c: Evolution of probability of household income decline

Figure 2d: Evolution of the inability to work as usual

Sources: HFPS and GMD: Q1 is poorest, Q5 is wealthiest.

Figure 2b suggests that approximately a quarter of the respondents in the average country stopped working since the beginning of the pandemic. Respondents are mainly household heads, however, and are therefore not representative of the full population of workers. Although we see a small increase in the share of the population in higher quintiles, a large share of respondents reported stopping work throughout the welfare distribution. Figure 2d compares the evolution of the impact of Covid-19 on the ability to work as usual for the bottom 40 and top 60 percent of the population. A larger share of the top 60 percent of the population were unable to work as usual in the early stages of the pandemic. The
magnitude of the difference ranged from five to ten percentage points during the spring and summer of 2020, and was statistically significant in the first five months. As was the case for total income declines, however, this distributional gap narrowed and had disappeared by October 2020.

Figure 3a shows the evolution of the difference between the bottom 40 and top 60 in those reporting that they had engaged in farming since the start of 2020. In the initial stages of the crisis, poor households were more likely to enter agriculture than wealthier households. Wealthier households also increased greatly farming activities later, in December and January, narrowing the gap back to pre-crisis levels. The shrinking differences in early 2021 are imprecisely estimated, however. Meanwhile, the share of households in the bottom 40 owning, even partially, a non-farm enterprise increased between July 2020 and March 2021 (Figure 3b). Like farming, wealthier households saw a large increase in non-farm enterprise ownership in early 2021. Overall, the results indicate that poor households were more likely to establish household enterprises in response to the crisis during the summer and fall of 2020.

Figure 3a: Evolution of share of households engaged in farming activities

Figure 3b: Evolution of share of households owning a non-farm enterprise

Food security

A major concern during Covid-19 has been its potential impact on food security of households in SSA. While the HFPS data contain several measures of food insecurity, we focus on the share of households in which an adult went without eating for a full day due to lack of money, because it is both objectively defined and an indicator of severe financial distress. As expected, poorer households report greater food insecurity than non-poor households. The difference is fairly modest, however, as about 17 percent of the bottom and second quintile report not eating for a fall day, while 11 percent of the top quintile does. This is consistent with the notion that the initial impacts of the crisis were widespread throughout the welfare distribution. Figure 4b presents how the share of households in the bottom 40 percent in which an adult went without eating for a day changed over the course of the pandemic, relative to households in the top 60 percent. Food insecurity increased significantly for the bottom 40 percent between April and June.

6 Results of the tests for statistical significance are not shown
2020, before falling dramatically between July and April 2021, suggesting a strong recovery for the poorest households in the second half of 2020.

Figure 4a: Share of households in which an adult went without eating for a whole day in the last month due to lack of resources (wave 1)

Figure 4b: Evolution of share of households in which an adult went without eating for a whole day in the last month due to lack of resources

---

**Children’s education:**

Perhaps the population that was most affected by Covid-19 is children, whose learning was adversely affected by school closures and the lack of alternative arrangements when the pandemic began. Figure 5a plots the share of households with children 6 to 18 that reported engaging in educational activities since school closures by predicted welfare quintile. There is a clear pattern of households in the bottom 3 consumption quintiles being about 15 percentage points less likely to report children engaged in educational activities than those in the top quintile. Figures 5b and 5c corroborate the notion that educational engagement was more difficult for poor households than wealthier households when it came to completing assignments provided by the teacher or watching education TV programs. Finally, Figure 5d plots how the share of children engaged in education activities since the beginning of the pandemic for the Bottom 40 and the Top 60 percent of the population. Engagement increased slightly for the top 60 between April and June, before falling dramatically between June and January. Engagement declined more gradually for the bottom 40, about 10 percentage points between April 2020 and January 2021. The difference between the bottom 40 and top 60 was statistically significant until November. This points to the widespread challenges faced by both wealthier and poorer households in continuing educational activities during such a prolonged crisis.

**Figure 5: The impact of Covid-19 on children’s educational engagement across the welfare distribution**
Safety nets and asset sales

One silver lining of the global pandemic has been that many governments were, at least initially, able to target assistance to poorer households. Figure 6a presents the share of households that reported receiving any form of government assistance in wave 1. Although only about 12 percent of households in
the lower two quintiles reported receiving assistance, the households in lower quintiles were significantly more likely to receive assistance compared to households in higher quintiles, with a difference of about 4 pp. However, figure 6c suggests that this progressive targeting was less apparent when considering data from all of waves, with a difference of about 1 percentage point. This difference changed little between April 2020 and January 2021.

Figures 6b and 6d examine the prevalence of asset sales in response to the crisis. Figure 5b indicates that poor households were substantially more likely to sell assets than wealthier ones in the immediate aftermath of the crisis, although less than 7 percent of respondents in the first and second quintile reported asset sales. Figure 6d suggests that over the course of the pandemic, poorer households initially became more likely to sell assets, after which asset sales declined for the poor during the summer of 2020. This is consistent with the decline in food insecurity for the poor reported above.
Conclusion

Understanding the impact of Covid-19 on households in different parts of the welfare distribution was an important factor when considering appropriate policy responses in response to the pandemic and will provide lessons for similar events in the future. We combined the High-Frequency Phone Surveys collected by national governments since the beginning of the pandemic with older surveys with information on household consumption to shed light on this question. We find that the initial impact of Covid-19 has been fairly widespread throughout the consumption distribution. In the initial aftermath of the crisis, poorer households were equally likely to report income declines but less likely to report the inability to work as usual. While government assistance slightly favored poorer households initially, this progressive targeting disappeared quickly and the share of households receiving government assistance was similar for the bottom 40 and top 60 percent for the year after the crisis. In the summer and fall of 2020, there was a large increase for households in the bottom 40 in farming activities and non-farm household enterprises. This coincided with economic recovery, as seen by large reductions in food insecurity and distressed asset sales among the bottom 40 during this time, as well as a reduction in the prevalence of income declines. However, the engagement of children in educational activities remains a concern. By the end of 2020, only about one third of households with children reported that children aged 6 to 18 were engaged in educational activities, a decline of about 15 percentage points.

To sum up, the results suggest that the initial impact of the shock was severe throughout the welfare distribution, and that there was a marked recovery for poor households associated with an increase in farming and non-farm household enterprises. However, poor households remained more likely than non-poor households to report income declines and not working as usual in the medium term. The impact on educational engagement was severe and felt throughout the distribution. The results highlight the importance of policies to ensure productivity in the agricultural sector, as well as non-agricultural household enterprises, in response to the crisis.

The analysis in this note could be extended in several directions. We do not closely examine variation in the impact of the global pandemic across the 14 countries in our analysis sample, or differences between urban and rural areas, which could provide further insights. We also do not examine the differences across various socio-economic characteristics such as age, gender, and labor market characteristics of individuals, although this is challenging in phone surveys because of the non-representative nature of the sample. As more face-to-face surveys are implemented and analyzed, a clearer picture will emerge of the distributional impacts of this prolonged pandemic in Sub-Saharan Africa.
References


Appendix

Appendix: S2S estimation model

Step 1: identify a set of variables that are available both in HFPS and GMD. We define the vector of regressors $X_1$ as the variables in Appendix table 1 plus the interaction terms. $X_1$ is common to all the countries.

**Appendix table 1: Level variables available in both GMD and HFPS**

<table>
<thead>
<tr>
<th>Variables</th>
<th>Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>hsize</td>
<td>Household size</td>
</tr>
<tr>
<td>hhsizeSq</td>
<td>Household size squared</td>
</tr>
<tr>
<td>children_share</td>
<td>Share of children residents in the household</td>
</tr>
<tr>
<td>elderly_share</td>
<td>Share of elderly residents in the household</td>
</tr>
<tr>
<td>urban</td>
<td>Area of residence (urban/rural)</td>
</tr>
<tr>
<td>departement</td>
<td>Subnational administrative area</td>
</tr>
<tr>
<td>phone</td>
<td>Phone Ownership (household level)</td>
</tr>
</tbody>
</table>

For each country,

Step 2: In the GMD dataset, we estimate $\log(welfare) = y^{GMD} = \beta_1 X_1^{GMD} + \varepsilon$ and use the “Lasso” method to select a set of regressors, $X_2^{GMD} \subset X_1^{GMD}$, we ensured that if an interaction term was selected by LASSO, the appropriate level terms were also included in the final model. For instance, if the interaction c.hsize#i.urban is selected by the “Lasso”, the variables hsize and urban should be in $X_2$.

Step 3: We estimate the model $\log(welfare) = y^{GMD} = \beta_2 X_2^{GMD} + \varepsilon$, ($X_2^{GMD}$ being allowed to differ across the countries)

Step 4: Predict welfare within the HFPS $\hat{y} = \hat{\beta}_2 X_2^{HFPS}$, using the coefficients of step 3.

**Appendix table 2: Adjusted R-squared of the S2S estimation model**

<table>
<thead>
<tr>
<th>Country</th>
<th>Adj. R-squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Burkina Faso</td>
<td>0.54</td>
</tr>
<tr>
<td>Ethiopia</td>
<td>0.36</td>
</tr>
<tr>
<td>Gabon</td>
<td>0.36</td>
</tr>
<tr>
<td>Ghana</td>
<td>0.43</td>
</tr>
<tr>
<td>Gambia</td>
<td>0.43</td>
</tr>
<tr>
<td>Mali</td>
<td>0.43</td>
</tr>
<tr>
<td>Mozambique</td>
<td>0.43</td>
</tr>
<tr>
<td>Malawi</td>
<td>0.53</td>
</tr>
<tr>
<td>Niger</td>
<td>0.46</td>
</tr>
<tr>
<td>Nigeria</td>
<td>0.55</td>
</tr>
<tr>
<td>Senegal</td>
<td>0.55</td>
</tr>
<tr>
<td>Sierra Leone</td>
<td>0.56</td>
</tr>
</tbody>
</table>
Appendix Table 3: Comparison of distribution of actual and predicted per capita household consumption
Appendix Figure 1: Comparison of variables between GMD and HFPS (wave 1) by country

Note: Percentage is used for non-numeric indicators.
Appendix Figure 2: Comparison of individual-level variables between GMD and HFPS (wave 1) by country

Note: Percentage is used for non numeric indicators

Appendix Figure 3: