Poverty projections and profiling from St Lucia High Frequency Phone Surveys using a SWIFT-COVID19 package

Chen, X., F. Fatima, U. Gray, Xueqi Li, S. C. Mathurin, A. L. Paffhausen, N. Yoshida, and K. Zhang

Abstract

This paper shows the results of poverty and inequality estimations using a SWIFT-COVID19 package into the Saint Lucia High-Frequency Phone Surveys (rounds 1 and 2). The SWIFT-COVID19 package includes the imputation of household expenditures using a SWIFT-Plus approach, a rapid poverty monitoring tool, and an adjustment for sampling weights to address a phone survey's sampling bias. The SWIFT-Plus approach could predict a sudden poverty increase in Afghanistan and an annual fluctuation of the poverty rate in Serbia. The package shows Saint Lucia likely experienced a sizeable increase in poverty between pre-COIVID and May 2020, with a slight reduction between May and August 2020. On the other hand, changes in inequality appear minimal. The poor are more likely to face food insecurity and job stoppages but more likely to have government assistance. The situation of the poor has also been changing. The situation of the poor in May was worse than that of the poor in August.

I. Introduction

The COVID-19 pandemic is spreading fast in Africa, and many countries in the region have implemented social distancing policies and lockdowns to contain the spread of COVID-19. However, as Ravallion (2020) has pointed out, these policies can have a high social cost, especially for the poor and vulnerable. The already poor are less likely to have existing buffers for times of crisis; they have little savings or food stocks, are heavily dependent on casual daily labor, and the majority cannot work from home. Given the significant impact that social distancing policies have had on in-person and casual daily labor, it is important to have reliable data to monitor these policies' benefits and costs, especially on the poor and the vulnerable.

The World Bank Group launched the COVID-19 High-Frequency Phone Surveys (COVID-19 HFPS) in April 2020 to monitor the socio-economic conditions of the COVID-19 pandemic. Phone surveys enable data collection even when enumerators cannot visit sample households due to social distancing policies and lockdowns, allowing for continual monitoring of poverty trends and the poor's socio-economic status. However, because ownership of phones is still limited among the poor in sub-Saharan Africa and other low-income countries, these surveys are likely biased toward non-poor and non-vulnerable groups. The phone surveys include proxies of poverty, such as questions regarding job losses, income changes, and food security. Still, these proxies sometimes present inconsistent results – e.g., employment status seems to be improving, but food security is not yet recovering.

Survey of Well-being via Instant and Frequent Tracking (SWIFT) is a rapid poverty monitoring tool. SWIFT can estimate household expenditures or incomes and related poverty statistics by collecting data on only 10 to 15 simple questions. Computer Assisted Personal Interview (CAPI) software is used to collect information on key variables, and the entire process takes 3-5 minutes. SWIFT develops a model to estimate household expenditures or incomes using data from the 10 to 15 questions on the most recent household survey and applying machine learning and multiple imputation techniques. By adding SWIFT questions in the COVID-19 HFPS, we can estimate household expenditure or income and poverty status of each household in the sample without largely increasing our interview time. Since most questions are simple yes-no questions, they are easy to integrate into phone interviews without heavily training enumerators. Estimating household expenditure and producing poverty statistics enables us to profile the poor and the non-poor separately and design pro-poor policies.

To overcome the challenges of the phone surveys discussed above, a new SWIFT package called SWIFT-COVID-19 was created. SWIFT-COVID-19 includes sampling weight adjustments to address non-poor biases in typical phone surveys, which includes propensity score weighting (originally proposed by Rosenbaum and Rubin (1983 and 1984)) and post-stratification weighting (such as raking or Stata's *maxentropy* command). The new package also adopts a new SWIFT methodology, called SWIFT Plus. Because the standard SWIFT approach tends to underestimate poverty during times of crisis, SWIFT Plus is used to produce accurate poverty rates during this time. More details will be discussed in the next section.

This note includes the following sections: Section 2 describes how pre-COVID and post-COVID poverty projections are produced using SWIFT and SWIFT Plus, Section 3 describes the weight adjustments, and Section 4 lists poverty profiles and poverty trends.

II. Poverty Projections from COVID-19 HFPS using the SWIFT-COVID-19 package

Survey of Well-being via Instant and Frequent Tracking (SWIFT) is used to estimate poverty rates from the COVID-19 HFPS. SWIFT combines machine learning techniques and the latest ICT technology to estimate household consumption expenditure and produce poverty statistics. SWIFT makes it possible for users to obtain reliable poverty data and profile the poor within budget. It collects only 10 to 15 questions on poverty correlates, such as ownership of assets, housing conditions, and household demographics; projects household income or expenditure using those correlates in a statistical model; and estimates

statistics on poverty and inequality from the projected income or expenditure data. SWIFT has proved its usefulness in over 50 countries on more than 100 projects.

Reliability of SWIFT in the COVID-19 pandemic

Supported by years of quality control efforts, SWIFT has produced reliable estimates on poverty, inequality, and income growth. Yoshida et al. (2020) include results of empirical tests on the reliability of the SWIFT methodology. SWIFT models are tested by using two rounds of comparable household expenditure data for a given country. The models are developed from the first round of data and applied to the second round to estimate poverty statistics. The poverty estimates are then compared with the official poverty rates to see how well the SWIFT estimates match up. As shown in Table 1, the differences between SWIFT estimates and the official poverty rates are small. All estimates are less than 1.5 percentage points away from official poverty rates, and in 5 out of 6 cases, the differences are statistically insignificant at the 5 percent level. The only exception is the estimation for Romania's rural area, where the estimate is slightly outside the 95% confidence interval. More evidence on the reliability of SWIFT estimates is available in Yoshida et al. (2020).

Table 1. SWIFT model prediction power over time

Country	year gap	Region	Absolute Difference
Uganda	**3**	Urban	**1.09%**
		Rural	**0.16%**
Romania	**1**	Urban	**0.03%**
		Rural	1.46%
Sri Lanka	**3**	Urban	**0.15%**
		Rural	**0.85%**

Note: Predictions are in bold lie within 95% confidence interval of original poverty rates.

However, Yoshida et al. (2020) found that SWIFT does not perform well during a large negative economic shock, like the COVID-19 pandemic. Afghanistan (2011 – 2016) and the West Bank and Gaza (2011 – 2016) both experienced severe economic downturns where the percentage of poor people increased by 16 and 14 percentage points, respectively. However, the standard SWIFT approach underestimated the poverty rate increases – estimating increases of only 5 and 6 percentage points in Afghanistan and the West Bank and Gaza, respectively.

Yoshida et al. (2020) show that underestimating a surge of poverty during economic downturns is due to the inclusion of slow-changing indicators, like asset ownership, in the standard SWIFT models. While asset ownership is highly correlated with household expenditure or income during times of stable economic growth, the correlation weakens during times of crisis when poverty surges. Due to the lack of active second-hand markets, households cannot easily sell many of their assets during a crisis, even when household income declines substantially. This leads to the standard SWIFT model producing underestimates of poverty during economic downturns.

Creation of SWIFT Plus

A modified approach, SWIFT Plus, was developed to overcome the standard SWIFT model's underestimation of poverty during severe economic downturns. While a standard SWIFT model selects indicators highly correlated with household expenditure or income, SWIFT Plus selects indicators that quickly reflect the economic conditions, even though they are only moderately correlated with household expenditure or income. Specifically, SWIFT Plus includes dummies for consumption of specific items like meat or shirts. Households tend to stop purchasing these items when their income declines, but resume

purchasing them once their income recovers. SWIFT Plus also includes economic sentiments, food security indicators, and employment conditions, all of which change quickly depending on the economic conditions. SWIFT Plus replaces time-invariant poverty correlates from the standard SWIFT model with the above-mentioned time-variant poverty correlates. Yoshida et al. (2020) provide evidence for SWIFT Plus. For both Afghanistan and the West Bank and Gaza cases, SWIFT Plus estimated substantial poverty increases which were very close to the actual increases.

The SWIFT Plus approach is adopted to estimate post-COVID poverty rates using the COVID-19 HFPS data. To run SWIFT Plus, time-variant indicators like consumption of specific items, food security, employment conditions, and economic sentiment are added into the COVID-19 HFPS questionnaire.

Pre-COVID poverty projections

To estimate the impact of the COVID-19 pandemic on poverty, both the pre-COVID and post-COVID poverty estimates are needed. If the latest household budget survey was conducted just before the COVID-19 outbreak, that survey's poverty rates can be treated as pre-COVID poverty rates. However, if the latest household budget survey was conducted even one year ago, the pre-COVID poverty rate could be different from the poverty rate estimated from that survey.

As discussed above, post-COVID poverty rates can be estimated using the SWIFT Plus approach replacing some time-invariant indicators with time-variant indicators collected via phone surveys. Because most countries did not collect data just prior to the COVID-19 outbreak, pre-COVID poverty rates must be estimated using data collected after the outbreak. The standard SWIFT model's utilization of time-invariant indicators serves well to estimate poverty rates just prior to the start of a crisis. For example, housing conditions cannot be changed unless households move to different houses, and such a move is unusual, especially during the beginning of a pandemic. Ownership of assets also do not change from the pre-COVID time unless households sell their assets, which can be difficult because many developing countries do not have active second-hand markets for consumer durables. Therefore, the current status of many indicators used in a standard SWIFT model can be used to estimate the pre-COVID poverty rates. Data on these time-invariant indicators can be collected alongside data on time-variant indicators in the first round of COVID-19 HFPS.

Table 3 illustrates how pre-COVID and post-COVID poverty rates can be estimated, showing an example of data from two rounds of COVID-19 HFPS. First, a standard SWIFT model (f_s) is estimated using a full set of time-invariant indicators (X) and a SWIFT Plus model (f_{sp}) is estimated using a subset of time-invariant indicators (X') and time-variant indicators (Z) from the latest household budget survey.

Ideally, to estimate pre-COVID poverty rates, we would insert data for the subset of time-invariant indicators ($\mathrm{X'}_0$) and time-variant indicators collected just prior to the COVID-19 outbreak (Z_0) into the SWIFT Plus model (f_{sp}); however, this data does not exist. Instead, we can use data for time-invariant indicators collected during round one of the COVID-19 HFPS (X_1), since status on these slow-moving indicators will likely be unchanged from the period of time just prior to the COVID-19 outbreak. We insert a full set of time-invariant indicators from round one COVID-19 HFPS into the standard SWIFT model (f_s) to estimate pre-COVID household expenditure (Y_0) and poverty rate (P_0).

We then insert round one COVID-19 HFPS time-invariant and variant variables ($\mathrm{X'}_1$ and Z_1) into the SWIFT Plus model (f_{sp}) to estimate household expenditures (Y_1) and poverty rate (P_1) for the round one time period. For round two estimations, we collect time-variant variables (Z_2) during round two of the COVID-19 HFPS. It is not necessary to collect round two time-invariant variables (X'_2) because they are very slow to change from the previous round (X'_1) (in other words, $X'_2 \approx X'_1$). Still, if there is room in round two's questionnaire, it is good to add questions on the time-invariant variables since they might also change, if only slowly. We estimate round two household expenditure and poverty rates by inserting both round two time-invariant and variant variables ($\mathrm{X'}_2$ and Z_2) into the SWIFT Plus model (f_{sp}).

Table 3. Illustration of SWIFT COVID-19 projections

	Pre-COVID (Round 0)	Round 1	Round 2
Time-invariant Variables: $\boldsymbol{X_0 = X_1 = X_2}$	X'_0 (unavailable)	X_1, X'_1	X'_2
Time-variant variables strongly correlated with real-time welfare level:	Z_0 (unavailable)	Z_1	Z_2
Household expenditures	$\boldsymbol{Y_0} = f_s(X_1)$	$\boldsymbol{Y_1} = f_{sp}(X'_1, Z_1)$	$\boldsymbol{Y_2} = f_{sp}(X'_2, Z_2)$
Poverty rates	$\boldsymbol{P_0}$	$\boldsymbol{P_1}$	$\boldsymbol{P_2}$

Note: X_t refers to a full set of time-invariant indicators in period t, and X'_t refers to a short set of time-invariant indicators in period t.

III. Reweighting to obtain nationally representative poverty estimates

One shortcoming of the COVID-19 HFPS is its lack of national representativeness in key statistics. People who respond to phone interviews may have systematically different characteristics as compared to people who do not respond to phone interviews. For example, in poor areas like rural Malawi, many poor households do not own phones, although most rich households do. The situation can also differ substantially between urban and rural areas. Even in a developing country like Malawi, telephone ownership in urban areas is close to 90 percent. Since phone ownership is essential for phone interviews, such an unbalanced distribution of phone ownership makes the collection of nationally representative data challenging. Besides the unbalanced phone ownership, responses to phone interviews are often not uniform. Like face-to-face interviews, rich households in urban areas tend to reject phone interviews more likely than poor households. As a result, data from phone surveys are unlikely to represent a country uniformly, and statistics from the data are often not nationally representative.

To address the possible limitations of a phone survey, we adjust sampling weights so that weighted averages of key statistics become nationally representative. The reweighting process has two major steps: (i) Propensity Score Weighting and (ii) Maxentropy or raking. This note briefs the steps.

Propensity Score Weighting (PSW)

Propensity Score Weighting (PSW) is designed to adjust a phone survey's sampling weights by comparing a nationally representative household survey, called a reference survey, with a phone survey. PSW merges the reference survey and the phone survey and estimates each household's probability in the merged data of being included in a phone survey. PSW then ranks all households in the merged data by the predicted probability and creates quintiles. The weights of households in the phone survey are adjusted so that each quintile's share of households in the reference survey becomes identical to that of the phone survey. More specifically, the weights of households in the phone survey are adjusted so that the sum of their weights in each quintile becomes identical to that of households in the reference survey.

To refine the weights further, we execute maxentropy or raking. Even after PSW, summary statistics in the phone survey could differ largely from those in the reference survey. Such differences can be real, particularly when a long time has passed between the reference and phone surveys. Still, it is unlikely that summary statistics of time-invariant (or slow-changing) indicators like household size, dependency ratios, household head's education attainment, or population shares of districts would change significantly.

Maxentropy and raking adjust weights to match the summary statistics of these time-invariant variables between the reference and phone survey. The following box briefly explains how maxentropy works.[1]

Box 1. Maxentropy

Maxentropy is a STATA command that selects weights that maximize entropy while matching averages of pre-selected indicators between the reference and phone surveys. The selection of indicators is important. The indicators need to be time-invariant or slow-changing. Otherwise, because there is some time between the reference and phone surveys, the averages of indicators can change. Ignoring the real changes and matching the averages between the two surveys can bias all statistics estimated from the phone survey. Therefore, it is important to select indicators that are time-invariant or slow-changing.[2] Indicators like household size, dependency ratio, and population shares of subnational units are such examples. However, since these indicators can also change over time and the speed of the change varies by country, it is always useful to look at trends of these indicators using the multiple rounds of comparable household surveys in the past before selecting the indicators for matching.

More details on all of the above-mentioned reweighting, Propensity Score Weighting, and maxentropy can be found in Zhang et al. (2021).

Application to the St Lucia COVID-19 High Frequency Phone Surveys

The aforementioned reweighting and poverty projection methods were applied to the St Lucia COVID-19 High Frequency Phone Surveys (round one and round two). Details are available in the annex.

IV. St Lucia COVID-19 High Frequency Phone Surveys (rounds one and two)

Round one of the St Lucia COVID-19 HFPS drew its sample from the database of telephone numbers from 179 out of 582 Enumeration Districts (EDs) spread across the eleven Saint Lucian regions. The database was collected as part of listing for another survey, a St Lucia Disaster Risk Management (DRM) & Poverty survey. The DRM & Poverty survey was supposed to be conducted via face-to-face interviews but was canceled due to the St Lucia government's social distancing policy. During the listing, all households in the selected EDs were asked to provide their telephone numbers, but the provision was left to each household's decision. Consequently, around 1,800 households provided their contact numbers during the listing exercise. The survey's final sample declined further to 1,093 phone numbers because some numbers were no longer active, and some households did not agree to participate in the HFPS. The final sample is spread across 141 EDs in the country.

The first round of HFPS was collected between May 5 and May 18, 2020. Although the sample for the original Saint Lucia DRM & Poverty survey was designed to be nationally representative, the final sample may not be as representative due to selection bias. To make all statistics from the data nationally representative, the sampling weights were adjusted using a propensity score weighting (PSW) technique and maxentropy, as mentioned above.

The second round of HFPS was collected between July 27 and August 20, 2020. All respondents from the first round of HFPS were contacted for interviews in the second round, but due to nonresponses, the second round's final sample is 900. Again, the sampling weights were adjusted. All statistics in this report were calculated using the adjusted weights to provide nationally representative results.

The St Lucia HFPS used the global core questionnaire prepared by the COVID-19 Questionnaire Working Group of the World Bank, with minor modifications, including questions for SWIFT poverty projections.

[1] Inputs for reweighitng process are available in annex 3.
[2] This identification of time invariant variables is also important when running SWIFT Plus.

V. Results – Poverty projections and profiles

The following section shows the trends in poverty and inequality from 2016 to August 2020, including estimates for the periods of time directly before and after the COVID-19 outbreak. This section also compares profiles on food security, employment status, and the coverage of social protection for several key groups: the pre-COVID poor, the poor in the first round of the HFPS, the poor in the second round of the HFPS, and the total national population. In the following discussion and figures, the "pre-COVID poor" refers to the group of individuals who would have been considered poor prior to the COVID-19 outbreak. Estimates for the pre-COVID poor were produced using the standard SWIFT methodology with time-invariant variables from the first round of COVID-19 HFPS. Estimates for the poor in the first and second round of COVID-19 HFPS were calculated using SWIFT Plus. The final SWIFT models are available in the annex (annex 2).

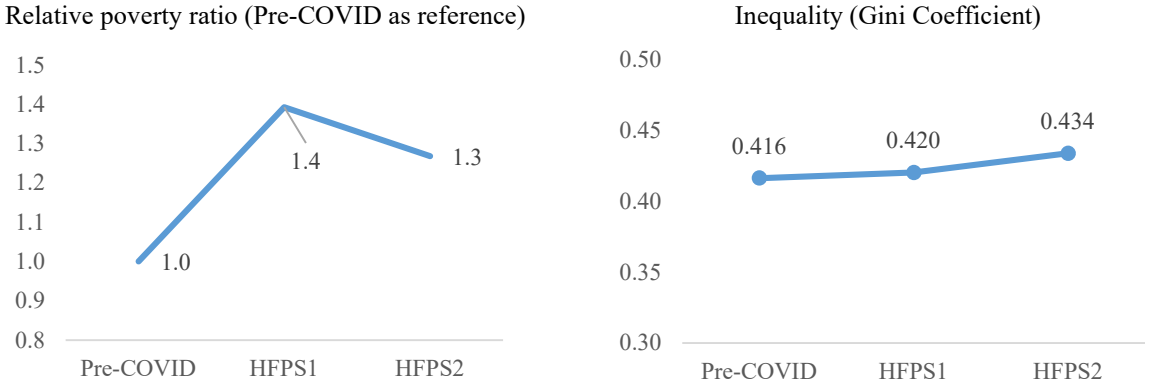
Poverty

Figure 1 shows the relative poverty score, a ratio of poverty rates with pre-COVID as reference. The poverty rates are estimated using the SWIFT Plus methodology. This relative poverty score has been used widely in the Latin America region to show the fluctuation of income poverty. According to this measure, poverty incidence has increased 40 percent (not 40 percentage points) between the pre-COVID era and May 2020 (round one COVID-19 HFPS). The incidence of poverty has then started to decline 10 percent from May to August 2020 (between COVID-19 HFPS rounds one and two). All the stated results use point estimates. To consider the margin of errors for this increase, we estimate a probability of having a poverty rate with a more than 5 percentage point increase. The probability was 85 percent between pre-COVID and May 2020. It declined significantly to 64 percent between May and August 2020.

Inequality

Based on the imputed consumption expenditures, the Gini coefficient can be estimated for pre-COVID, round one, and round two data. All of the Gini coefficients are above 0.4, indicating a sizeable income gap in St Lucia, even before the COVID-19 outbreak. After the COVID-19 outbreak, the Gini coefficient has been increasing slightly to 0.42 in May 2020 (round one) and then to 0.434 in August 2020 (round two). The upward trend shows that the recovery from the COVID-19 pandemic is not uniform and can increase income inequality further.

Figure 1. Trends between HBS 2016 and HFPS Round 2 (August 2020)



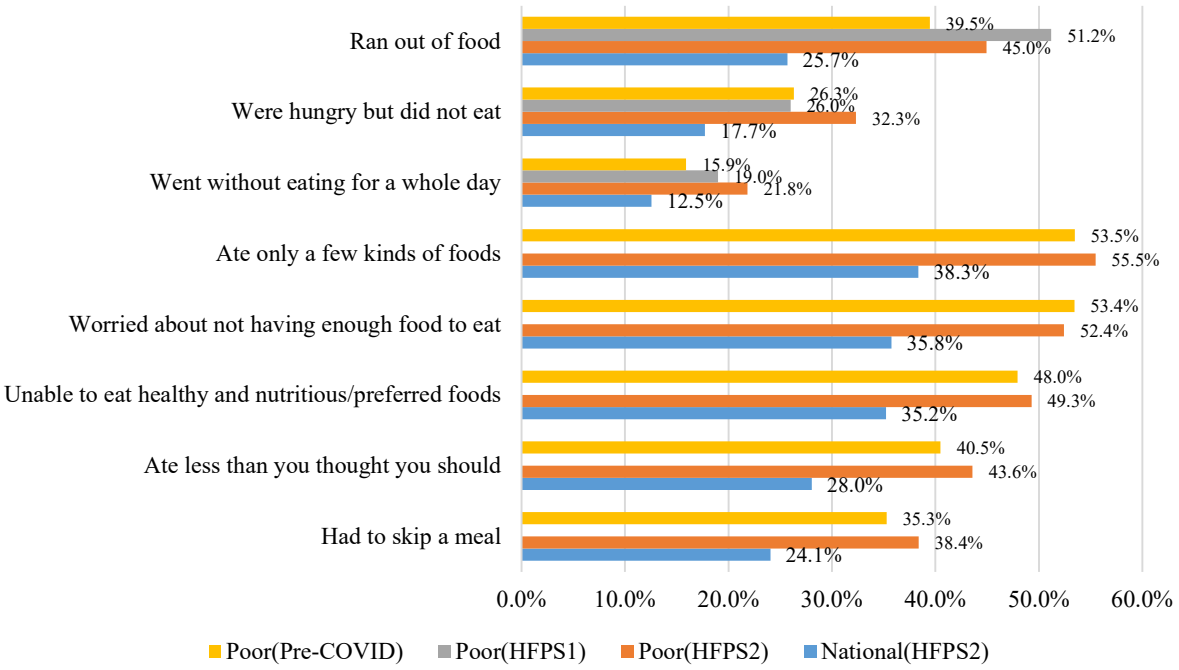
Source: Authors' estimation using data from HBS 2016 and HFPS round 1 and round 2

Food security

Food security is measured by eight questions included in the Food and Agriculture Organization (FAO)'s food security measurement. Round one included only three questions, while round two included all eight questions. The surveys yielded two major findings. First, the poor's food security conditions were far worse than the national average. For example, in almost one-third of poor households surveyed in round two, at least one adult member experienced hunger but did not eat anything during one or more days in the last 30 days, as compared to only eighteen percent of all households. Second, the poor's food security has worsened severely since the COVID-19 outbreak. However, the trend from round one is mixed — two out of three indicators show a significant deterioration in food security, but the rest shows some improvement.

Figure 3. Comparison of food security indicators across different groups

During the last 30 days, was there a time when any member in the household...

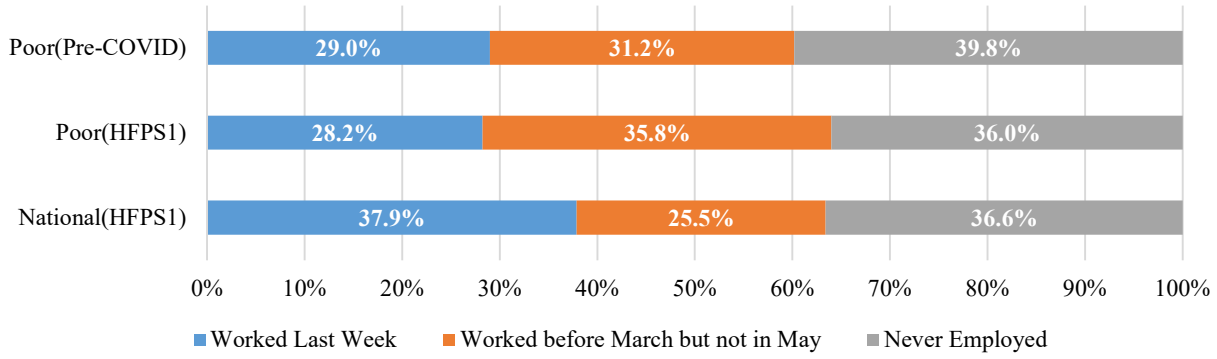


Source: Authors' estimation using data from HBS 2016 and HFPS round 1 and round 2

Employment status

In round one, all respondents were asked whether they were working last week and, if not, whether they were working before the COVID-19 outbreak (i.e., before March 20th, 2020). According to the round one data, only 29 percent of the pre-COVID poor were working in May 2020 (during the round one HFPS) and 31 percent reported to have worked prior to the COVID-19 outbreak but not in May 2020. The percentage of job stoppage among the pre-COVID poor is significantly higher than the national average (31 versus 26 percent). The poor in round one faced an even worse situation – with their percentage of job stoppage at 36 percent.

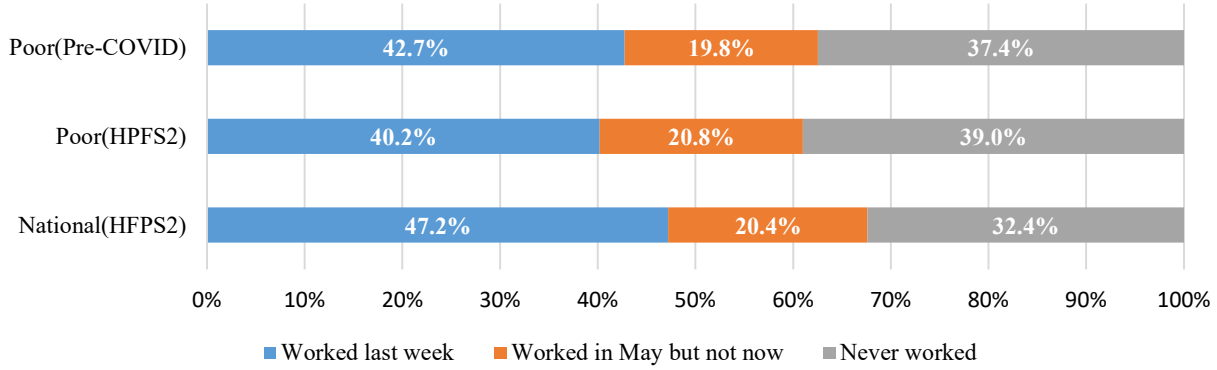
Figure 4. Comparison of employment status across different groups, as of HFPS round one (May 2020)



Source: Authors' estimation using data from HBS 2016 and HFPS round 1 and round 2

Employment status appears to have improved between rounds one and two of the COVID-19 HFPS. However, the improvement is not uniform. Around 43 percent of the pre-COVID poor worked in August 2020 (during the round two HFPS), compared to 29 percent in May 2020. However, the employment status of the pre-COVID poor still remained worse than the national average (43 versus 47 percent). The poor in August 2020 faced an even lower employment rate than the pre-COVID poor.

Figure 5. Comparison of employment status across different groups, as of HFPS round two (August 2020)

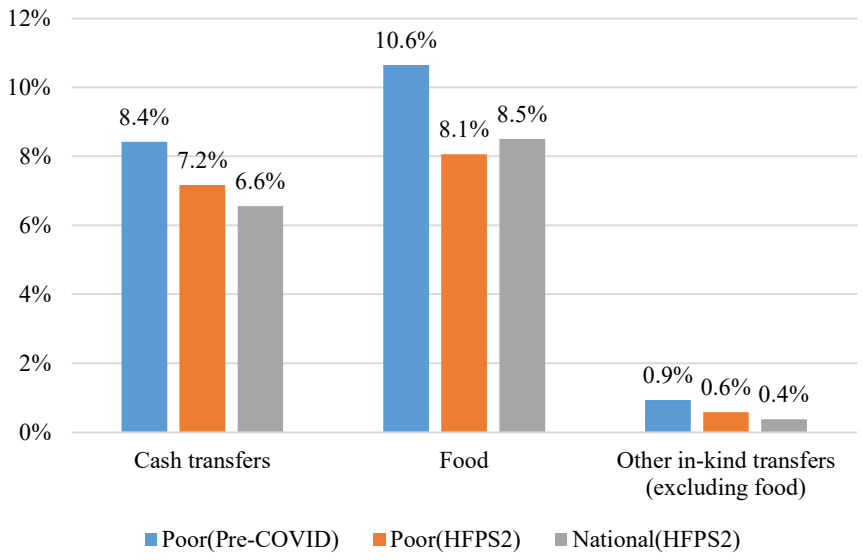


Source: Authors' estimation using data from HBS 2016 and HFPS round 1 and round 2

Coverage of government assistance

In St Lucia, the coverage of government assistance was pro-poor in that those who were already poor before the COVID-19 outbreak was higher than that of the national average population. However, the government assistance is less pro-poor against the poor in round 2. The coverages of cash transfer and other in-kind assistance to the poor in the round 2 are higher than those of the national average, but the coverage of in-kind assistance to the poor in round 2 is less than that of the national average. Note that the round 2 data were collected before the government expanded the social assistance program. It will be useful to estimate the coverage of the poor and the national average population in the newly expanded program to see the pro-poorness of it.

Figure 6. Comparison of government assistance across different groups



Source: Authors' estimation using data from HBS 2016 and HFPS round 1 and round 2

VI. Conclusion

The COVID-19 pandemic has affected how to collect data and how to estimate poverty and inequality. Since the beginning of the pandemic, St Lucia has collected two rounds of COVID-19 High-Frequency Phone Surveys (COVID-19 HFPS), the first in May and the second in August of 2020, via phone interview. Poverty incidence and inequality have been estimated from this data using the SWIFT-COVID-19 package, which adjusts the original SWIFT methodology to be more responsive to sudden economic downturns and addresses sampling bias due to phone interviews by reweighting. This note includes poverty and inequality estimates for the periods before the COVID-19 outbreak and the two rounds of HFPS (May and August 2020).

Estimates show a spike in poverty in May 2020 compared to poverty rates before the COVID-19 outbreak – the relative poverty score increased from 1 to 1.4 and the probability of poverty increasing more than 5 percent was 85 percent. This increase in poverty is a sharp reversal of the trend before the COVID-19 outbreak. Subsequently, poverty has decreased slightly between May and August 2020. Inequality did not change much in May but started to increase slightly between May and August.

On most indicators, the poor have experienced increased food insecurity and fared worse than the national average during the pandemic. The poor have also experienced worse employment conditions, with higher rates of job stoppage, than the national average, though this has improved slightly between May and August of 2020 even among the poor. Lastly, government assistance remans pro-poor but less so for the poor in the round 2. Since the round 2 data were collected before the government expanded the social assistance program. It will be useful to estimate the coverage of the poor and the national average population in the newly expanded program to see the pro-poorness of it.

References

Rosenbaum, P. R., and D. B. Rubin. (1983). "The Central Role of the Propensity Score in Observational Studies for Casual Effects." *Biometrika* 70 (1): 41-55.

Rosenbaum, P.R., and D.B. Rubin. (1984). "Reducing Bias in Observational Studies using Subclassification on the Propensity Score." *Journal of the American Statistical Association*. 79: 516-524.

Yoshida, N., X. Chen, S. Takamatsu, K. Yoshimura, S. Malgioglio, and S. Shivakumaran. (2020) "The Concept and Empirical Evidence of SWIFT Methodology." Mimeo.

Zhang, K., Chen, X., and Yoshida, N. (2021). "Phone Survey Reweighting: Case Studies from the Philippines, St. Lucia, Zimbabwe, and Uganda." *World Bank Group*. 1-34.

Annex 1. Estimation of poverty trends

Model	Timing	Mean	Std. Err.	95% Confidence Interval	
Pre-COVID Model	HBS2016	26.86%	1.68%	23.57%	30.15%
	Pre-COIVD	23.86%	2.65%	18.62%	29.10%
Post-COVID Model	HFPS_Round1	33.25%	3.41%	26.40%	40.11%
	HFPS_Round2	30.33%	3.33%	23.66%	37.01%

Annex 2. Models for estimating poverty rates for Pre-COVID and Post-COVID.

Table 3.1 Pre-COVID Model (St. Lucia)

Linear Regression						Number of obs =		1487
						F(22,1464) =		54.23
						Prob > F =		0
						R-squared =		0.5549
						Root MSE =		0.49729
lnexp	Coef.	Std. Err.	t	P > t	[95% Confidence Interval]		HBS2016	HFPS1
urban	0.111073	0.035825	3.1	0.002	0.040799	0.181347	0.7	0.71
hhsize	-0.26489	0.015542	-17.04	0	-0.29538	-0.2344	3.07	3.06
hhsize2	1.129226	0.115679	9.76	0	0.902313	1.35614	0.13	0.13
hh_sex	0.12295	0.034118	3.6	0	0.056024	0.189875	0.6	0.6
Freezer	0.169498	0.046198	3.67	0	0.078876	0.260119	0.13	0.25
drinking2	0.208881	0.036974	5.65	0	0.136353	0.281409	0.46	0.61
depend	-0.2335	0.057815	-4.04	0	-0.34691	-0.12009	0.33	0.33
Washing_machine	0.203321	0.03707	5.48	0	0.130605	0.276037	0.56	0.64
tank	0.155367	0.040886	3.8	0	0.075166	0.235568	0.24	0.38
tv	0.125073	0.049406	2.53	0.011	0.028159	0.221987	0.82	0.75
wall1	-0.23767	0.053782	-4.42	0	-0.34317	-0.13217	0.13	0.1
Elecgas_stove	-0.22257	0.101749	-2.19	0.029	-0.42216	-0.02298	0.91	0.97
wall3	-0.13997	0.036979	-3.79	0	-0.21251	-0.06744	0.18	0.2
land_owner1	0.126037	0.034408	3.66	0	0.058543	0.193531	0.29	0.31
drinking3	0.119778	0.081537	1.47	0.142	-0.04016	0.27972	0.08	0.08
Jewellery	0.143275	0.042041	3.41	0.001	0.060809	0.225742	0.14	0.25
Refrigerator	0.134878	0.059329	2.27	0.023	0.0185	0.251256	0.81	0.87
heater	0.283024	0.080701	3.51	0	0.124721	0.441326	0.07	0.14
fuel	0.454481	0.113563	4	0	0.231717	0.677245	0.94	0.97
lawn_mower	0.261132	0.082803	3.15	0.002	0.098707	0.423558	0.08	0.17
toilet2	-0.21071	0.052469	-4.02	0	-0.31363	-0.10779	0.19	0.18
drinking5	0.277667	0.05042	5.51	0	0.178764	0.376571	0.16	0.17
_cons	9.18843	0.111277	82.57	0	8.970151	9.40671		

Table 3.2: Pre-COVID Poverty Estimation

	Mean	Std. Err.	[95% Conf. Interval]	
HBS2016	26.86%	1.68%	23.57%	30.15%
HFPS1_simulated	24.04%	3.15%	17.73%	30.35%

(1) The definition of drinking2, drinking3, drinking5 are slightly modified in the 2nd round analysis.

Table 3.3 Post-COVID Model (St. Lucia)

Linear regression

Number of obs	=	1474
F(13, 1460)	=	40.06
Prob > F	=	0
R-squared	=	0.351
Root MSE	=	0.5969

lnexp	Coef.	Std. Err.	t	P>t	[95% Confidence Interval]		HBS2016	HFPS1	HFPS2
district1	-0.22483	0.061936	-3.63	0	-0.34633	-0.10334	0.15	0.18	0.18
district2	-0.15473	0.057811	-2.68	0.008	-0.26813	-0.04133	0.25	0.24	0.23
district3	-0.18208	0.075515	-2.41	0.016	-0.33021	-0.03395	0.05	0.05	0.04
depend	-0.31319	0.077066	-4.06	0	-0.46437	-0.16202	0.33	0.33	0.33
district7	-0.31505	0.074561	-4.23	0	-0.4613	-0.16879	0.09	0.08	0.08
urban	0.297594	0.058603	5.08	0	0.182638	0.41255	0.7	0.71	0.7
hhsize	-0.21606	0.021157	-10.21	0	-0.25756	-0.17456	3.07	3.06	3.07
hhsize2	0.839599	0.19011	4.42	0	0.466681	1.212517	0.13	0.13	0.13
hh_sex	0.105181	0.037934	2.77	0.006	0.03077	0.179591	0.6	0.6	0.59
hh_age	0.004675	0.001331	3.51	0	0.002064	0.007285	52.26	54.76	54.69
FIES_8	-0.22545	0.09148	-2.46	0.014	-0.4049	-0.04601	0.07	0.1	0.13
FIES_6	-0.37788	0.050399	-7.5	0	-0.47675	-0.27902	0.2	0.3	0.26
employ	0.16493	0.043833	3.76	0	0.078948	0.250913	0.65	0.38	0.48
_cons	9.614381	0.103344	93.03	0	9.411662	9.817101			

Table 3.4 Post-COVID Poverty Estimation

	Mean	Std. Err.	[95% Confidence Interval]	
HBS2016	26.86%	1.68%	23.57%	30.15%
HFPS1_simulated	33.25%	3.41%	26.40%	40.11%
HBS2016	26.86%	1.66%	23.60%	30.12%
HFPS2_simulated	30.33%	3.33%	23.66%	37.01%

(1) This model excludes all housing variables as they are time-invariant.
(2) This model excludes all consumption dummies due to inconsistant way of asking the questions compared to HBS2016.

Table 3.5 Poverty Estimation

		Mean	Std. Err.	[95% Confidence Interval	
Pre-COVID Model	HBS2016	26.86%	1.68%	23.57%	30.15%
	Pre-COVID_simulated	24.04%	3.15%	17.73%	30.35%
Post-COVID Model	HBS2016	26.86%	1.68%	23.57%	30.15%
	HFPS_Round1_simulated	33.25%	3.41%	26.40%	40.11%
	HBS2016	26.86%	1.66%	23.60%	30.12%
	HFPS_Round2_simulated	30.33%	3.33%	23.66%	37.01%

(1) The 3 poverty estimations for different time period are genereated based on 2 rounds of HFPS data collection. The poverty estimation for pre-COVID and Round 1 are generated using round 1 data, but with different models. The poverty estimation for Round 2 are generated using round 2 data. (2) The poverty estimation are consistant between round 1 and round 2 as they are based on the same model.

Annex 3. Inputs for reweighting

- Surveys Needed o 1) **Reference survey that is nationally representative**. For example, in the Philippines study, we used the survey conducted in 2018, i.e., Family Income and Expenditure Survey

2018 (FIES 2018) o 2) **High**

Frequency Phone Survey

- Key Variables Needed: **time-invariant variables (listed below)** o Our goal is to make the phone survey resemble the distribution of the nationally representative survey as much as possible. To achieve this goal, we need to compare variables that are time-invariant between the two surveys. If these variables are close enough across the two surveys, we can safely conclude that the phone survey has resembled the reference survey quite well, or, the reweighting has been implemented successfully.

o Take the Philippines reweighting procedure as an example - we used the following time-invariant variables as targets to be matched across surveys:

- **household size**
- **household size squared**
- **dependent share**
- **urban/rural shares**
- **district-level population totals**
- **highest educational attainment of household heads**
- **the age of household head** o From both surveys, we also need the initial weights created before data collection. These weights serve as a starting point for weight adjustment (household weight (variable name: **WT),** population weight (variable name: **popweight)**).