Methodological Note on The Gambia High Frequency Phone Survey (HFPS)[1]

December 2021

Executive Summary

The objective of this note is to document the methodology utilized for the Gambia High Frequency Phone Survey (HFPS, 2020/21). The objective of the HFPS was to rapidly monitor the impact of the COVID-19 pandemic on households across the country. The survey comprised of 10 rounds of data collection between August 2020 and December 2021. A total of 1500 households were sampled from the 2018 Gambia Labor Force Survey (GLFS) and were contacted (in some cases without success) in each round of the survey. The sample is stratified at three levels: the capital city area-i.e the Banjul and Kanifing area; other urban area; and rural area. The data collected is representative at national and strata level. In each household, the most knowledgeable household member (mostly the household head) is interviewed over the phone.

The survey collected data on nineteen different topics across the ten waves. The topics includes demographic details about members of the household- detailed household roster data, labor market activities, knowledge about COVID-19, income loss, access to services, household wellbeing, food security, social cohesion, coping and social safety, private sector, education, remittances and social assistance, migration and remittances, housing, vaccine, disability, poverty, COVID-19 effects on children, climate events and agriculture. Certain modules such as household demographics, COVID-19 awareness and labor market activities were considered core modules and included in nearly every round of the survey.

Attrition from the HFPS is relatively low with those leaving the sample selective among some dimensions. In the first wave (August 2020), about 1,437 households were successfully surveyed and in the latest wave (December 2021, 18 months after first wave), 886 households were surveyed. This suggests an overall attrition rate of 38 percent, averaging 16 percent per round of survey. The last wave of the survey coincided with the 2021 Presidential elections. Additionally, the duration for data collection was shortened. These reasons may in part be responsible for the high attrition during this wave. Overall, the main reason reported for attrition include phone turned off (more than half), nobody answering and refusal. Attrition is correlated with gender, poverty status, and strata of residence of the respondent. The results suggest that males are less likely to miss at least one of the waves. Furthermore, those coming from the poorest 40 percent households and those in the rural areas are significantly more likely to participate in all rounds of the survey. This may in part be due to the incentive provided to respondents after every round of data collection- respondents receive airtime worth GMD 80 as an incentive for participating in the survey. Poorer households are more likely to continue to participate due to the incentive provided.

1. Sample design and implementation

The aim of the High Frequency Phone Survey was to collect data to enable the rapid monitoring of the COVID-19 impact on households in The Gambia. This survey is conducted in The Gambia on a subsample of the 2018 Gambia Labor Force Survey (GLFS) conducted between November 2017 and July 2018 by the Gambia Bureau of Statistics (GBoS). The GLFS dataset has 5,987 households of which 5,531 (92 percent) have valid phone numbers.[2] All households with

---

[2] The valid numbers starting with 2,3,6,7 or 9 and have a total of 7 digits.

valid phone numbers formed the population of households sampled for the HFPS. Data collection for the HFPS was conducted in ten-rounds and on a panel of about 1500 households. The first round of data collection was done in August 2020 using a sample of 1557 households, of which 1437 interviews were completed. Throughout the survey, the sample of households were stratified at three levels: the capital city areas- i.e Banjul and Kanifing area, other urban area and the rural area. By using population weights, data from the survey is representative at national and strata levels.

Overall, households in the HFPS sample were quite similar with the non-sampled housheolds in the base sample- i.e the 2018 labor foce survey; with few nuances. Notably, households in the HFPS sample residing in the Kanifing and Banjul strata are slightly larger. Similarly, for other urban areas, households in the HFPS sample are also larger and they are also relatively richer. Among the rural households, the HFPS sample is more likely to have household head who attended school. Furthermore, households from the sample without a valid phone number are selective. Those with valid phone numbers are relatively richer and they are more likely to have migrants. Finally, the survey respondents are statistically different from face-to-face would be respondents (household heads). The heads are statistically more likely to be males, older, more likely to be without formal education, and less likely to achieve tertiary education.

In each household, the most knowledgeable member (mostly the household head) is surveyed and provided an incentive in the form of airtime/mobile credit. Households were contacted via phone calls. Interviewers make several attempts to reach households throughout the data collection period. Additionally, information about best times to call were obtained in each wave to ease contacting households. In first wave of the survey, 1,437 households were successfully surveyed which represents an impressive 96 percent of the targeted 1557 households. In each wave, participating households receive incentives for participation in the form of airtime/mobile credit amounting to 80 GMD (about 1.50$). Overall, the high frequency phone survey covered 10 waves from August 2020 to December 2021.

Population weights were computed (by adjusting the weights from the baseline labor force survey) to ensure representativeness of the data. The weights used in the phone surveys are an adjustment of household weights in the baseline labor force survey. This adjustment is based on the calculation of the probability of a household responding to the phone survey which is computed from two components: the selection of the household to participate in the HFPS survey; and whether the household in question completes the survey in a particular round of data collection (i.e answers the questions in all the modules of the questionnaire administered). The response probability is computed by modeling the response/non-response on a set of demographic variables of the household (household size, dependency ratio, standard of living of the household measured by an asset index, age, gender, marital status, level of education, employment status, and sector of occupation of the head). To capture the likely dissimilarities between the different geographic strata, categorized into the Banjul and Kanifing, the other urban areas and the rural areas, a by-strata model is compiled. In each strata a stepwise modeling is conducted to select the above-mentioned variables that are most relevant in the particular context of the strata. The significance level for removal from the model is fixed to 15% and significance level for addition to the model is fixed to 5%. Prior weights are then obtained by multiplying the initial weight (baseline survey) by the inverse of the response probability. Finally, to preserve the structure of the total population and the distribution of households within strata, the entropy maximization algorithm is applied to compute the final weights used in the rapid surveys' analysis.

2. Description of topics covered by round of survey

The objective of the HFPS is to monitor the impact of the COVID-19 pandemic on household welfare. Data was collected on a series of topics which will summarized in this section. The questionnaire of a typical round of the survey comprised of a core section of questions which are typically included in every round of data collection; and a section of rotating modules- often based on events and emerging data needs. The core modules often include questions on household demographics, labor market activities and COVID-19 awareness. Table 1 below shows the various topics which were covered in the survey.

The core modules which are typically included in each round of the survey included questions on household composition/ roster, COVID-19 awareness and labor market activities. On household composition, the household roster data from the LFS 2018 was used as a baseline. In the subsequent rounds of data collection, respondents were asked to give an update of whether all previously listed household members were still part of the household and if they are no longer part of the household, the reason for leaving the household was recorded. Furthermore, if another individual has joined the household, their reason for joining, relationship with household head, sex, and age were recorded. Similarly, in the labor market activities section, questions on employment were included and elicited in all rounds of data collection with the data from 2018 LFS as the baseline. The employment section collected information on employment status, reasons for unemployment, working from home, sector of employment, transitions into and out of employment and sectors. In the first five waves, these outcomes were collected only from the main respondent of the survey (mostly the household head). However in the subsequent five waves, the outcomes of all household members aged 7 years and above were elicited through proxy reporting of the main respondent.

Questions about income and knowledge about COVID-19 were also frequent in the survey. The COVID-19 module was introduced in eight waves (all but the second and third wave) while the income topic was covered in seven waves (all but the fourth, fifth, and seventh wave). In terms of outcomes, the COVID-19 module collected information on awareness about the virus, sources of information about the virus, main prevention methods adopted, knowledge about vaccines, willingness to be vaccinated, vaccination status (type of vaccine shot – single/double), and reasons for taking or not taking up the vaccine; and measures and knowledge of government response taken to reduce the spread of the virus. The income module includes outcomes on the main sources of household income, income change (increased, stayed the same or reduced compared to pre-pandemic era), and main reason for the observed change in income (COVID-19, remittances, seasonality of work, social assistance).

Information on access to services, household wellbeing and food security were also collected. The module on access to servises is divided into different sub-topics – access to markets (household being able to buy and stock up on basic commodities and reasons for not if any), access to health services (household member visited medical care centers, reasons for visit, access to medical care, and reasons for not if any, and access to health insurance), access to educational services (educational activities during school closures, type of communication with teachers, and school resumptions), access to financial services (usage and access of different financial operations, and reasons for failure to access any), access to transportation services, and access to internet services.

On household wellbeing, the elicited outcomes include the welfare status of the household and comparison with pre-COVID-19 era, subjective household wellbeing scale (rich, medium, poor, very poor), expected 12 months wellbeing scale, concerned about the effects of COVID-19 on

household wellbeing[3], rating of government social and mitigating response to the COVID-19 pandemic.

Finally, the food security section elicited information on whether any household member is worried about not eating enough, unable to eat healthy and nutritious food, ate only a few kinds of foods, skip meals, ran out of food, hungry but unable to eat, and spend the whole day without eating. These questions were identified based on FAO's Food Insecurity Experience Scale (FIES) methodology to construct indicators of moderate and severe food insecurity.[4]

Table 1: Topics Covered by the High Frequency Phone Survey, by Waves

		WAVES									
		Wave 1	Wave 2	Wave 3	Wave 4	Wave 5	Wave 6	Wave 7	Wave 8	Wave 9	Wave 10
		Aug-20	Oct-20	Dec-20	Feb-21	Apr-21	Jun-21	Jul-21	Aug-21	Oct-21	Dec-21
No	Topic										
1	Household Composition/Roster	√	√	√	√	√	√	√	√	√	√
2	Knowledge about COVID-19	√			√	√	√	√	√	√	√
3	Access to Services	√	√	√					√		
4	Employment	√	√	√	√	√	√	√	√	√	√
5	Income	√	√	√			√		√	√	√
6	Wellbeing	√	√				√				√
7	Coping and Safety	√					√				
8	Food security		√				√			√	√
9	Private sector		√							√	
10	Housing			√							
11	Vaccine			√							
12	Social cohesion			√		√				√	
13	Education				√			√			
14	Remittances and Social Assistance				√			√			
15	Disability					√					
16	Poverty					√					
17	Migration and Remittances						√				√
18	Effects of COVID-19 and Children							√			
19	Climate Event and Agriculture								√		

Source: HFPS Questionnaire, 2020/21.

The HFPS also included topics on social cohesion, household coping and safety strategy, private sector, education, social assistances, and migration and remittances. Outcomes

---

[3] A module was included to collect data on strong correlates of household wellbeing identified from the 2015 Integrated Household Survey (IHS). This data was used to construct a profile of the poor based on a technique developed by the World Bank called SWIFT- see https://www.ifc.org/wps/wcm/connect/64f11adb-ab01-4207-93cd-dd2cc51af16c/SWIFT-booklet-05.pdf?MOD=AJPERES&CVID=m9Or9Ia.

[4] See https://www.fao.org/policy-support/tools-and-publications/resources-details/en/c/1236494/

included in the social cohesion module include fear of violence in neighborhood, community conflicts, service delivery, and social contract. The coping and safety module collected information on negative shocks affecting households and the main coping mechanisms employed by the household while the private sector module include outcomes on family business of the household, sector of the business, status and revenue of the business, and challenges faced by the household. Furthermore, the education module elicited information on children education including registration, educational activities conducted during school closures, and steps taken by parents, schools and government to prevent the spread of COVID-19. The migration, social assistance and remittances modules elicited outcomes on whether households have internal, international and/or returned migrants, whether households received any local or international transfers, monetary value and sources of the transfers, channels (and whether they have changed since COVID-19) of the transfers, social transfers (cash, food aid, other in kind).

Finally, modules on housing, vaccine, disability, effects of COVID-19 on children, and climate event and agriculture were included in at least one of the waves of the HFPS. The housing module in wave 3 included information residential status of the household (rent or own), value of rent, number of rooms, residency duration, reason and associated challenges of moving into the residence, while the vaccine module also collected in wave 3 has information on opinions about vaccination, willingness to take up, and willingness to pay for the vaccination. The disability and poverty modules all collected in wave 5 included questions on whether household has Persons With Disabilities (PWDs), how the pandemic affected their livelihood, difficulty in accessing their routine medical care, and support received and wished to receive. Similarly, modules on effects of COVID-19 on children, and climate event and agriculture were included in waves 7 and 8, respectively. The effects of COVID-19 includes the negative/positive effects of COVID-19 on children, support received by children, best support needed by children, challenges in protecting children from COVID-19, main places of risks for children, and policy gaps in the response of COVID-19 to mitigate its impact on children. On the other hand, the climate events and agriculture module have information on whether and how household was negatively affected by 7th/8th July 2021 windstorm or other negative shocks, coping mechanisms and monetary value of mitigating the shocks, opinion on the normality of the rainy season, harvest expectations, challenges of the farming season, and agricultural support needed from government and development partners.

3. Profile and selection of surveyed households and respondents

In this section, we describe the profile of households and respondents who participated in HFPS and do a comparison with households who were not sampled and those without valid mobile phone numbers. Recall that the HFPS sample is drawn from 5,531 households from the LFS, 2018 sample with valid phone numbers. Out of the total 5,987 households in the LFS, 1,496 households were successfully sampled and surveyed in at least one of the 10 rounds. Therefore, questions remain on what type of households or individuals made the HFPS sample and how selective (or representative) is the HFPS compared to the LFS, 2018 sample.

Table 2 below shows summary statistics of the household characteristics of households belonging to the HFPS and those households not part of the HFPS. We provide the descriptive statistics by grouping households into the three different strata: Banjul and Kanifing, other urban areas, and rural areas. The results suggest that the HFPS sample households are quite similar to those not part of the sample with few nuances. Notably, households in the HFPS sample residing in the Banjul/Kanifing area are slightly larger (7.5 vs 6.4 people). Similarly, for other urban areas,

households in the HFPS sample are also larger (9.6 vs 8 people) and they are also relatively richer as indicated by the average asset index (0.57 vs 0.55). Finally, among the rural households, the HFPS sample is more likely to have household head who attended school.

Table 2: Summary Statistics: Statistical Differences between HFPS and Non-HFPS

|	(1) HFPS Sample Mean	(2) Non-HFS Sample Mean	(3) T-test (1-2) p-value
Panel A: Banjul and Kanifing strata (household head/household characteristics)			
Age	45.862	46.303	0.602
Female	0.294	0.247	0.218
Household size	7.516	6.424	0.016*
Asset index	0.631	0.620	0.440
Working-age population	0.891	0.861	0.211
Employment status	0.720	0.725	0.929
HH has external migrants	0.205	0.154	0.054
Has attended school	0.993	0.998	0.293
Migrated internally	0.558	0.501	0.099
Unemployed	0.042	0.034	0.228
Observations	431	840	1291
Panel B: Other urban areas strata (household head/household characteristics)			
Age	47.685	47.090	0.942
Female	0.146	0.208	0.087
Household size	9.552	8.013	0.008**
Asset index	0.573	0.545	0.000***
Working-age population	0.885	0.888	0.542
Employment status	0.743	0.692	0.080
HH has external migrants	0.151	0.164	0.955
Has attended school	0.999	0.993	0.692
Migrated internally	0.454	0.426	0.217
Unemployed	0.077	0.106	0.137
Observations	530	1040	1570
Panel C: Rural areas strata (household head/household characteristics)			
Age	49.103	49.956	0.444
Female	0.147	0.157	0.463
Household size	11.030	10.475	0.101
Asset index	0.484	0.465	0.169
Working-age population	0.861	0.826	0.411
Employment status	0.527	0.541	0.076
HH has external migrants	0.136	0.163	0.782
Has attended school	1.000	0.996	0.045*
Migrated internally	0.232	0.228	0.187
Unemployed	0.261	0.230	0.856
Observations	509	2549	3058

Notes: HFPS sample represents households or individuals surveyed in the high frequency survey while LFS sample are those who are not part of this sample but surveyed in LFS 2018. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. There are 26 households from HFPS and 62 Non-HFPS household that are omitted due to missing asset index.
Source: HFPS 2020/21 and The Gambia Labor Force Survey, 2018

A few differences exist between households in the HFPS sample and those in the LFS but not part of the sample frame- i.e households in the LFS without a valid phone number. Households in the HFPS sample are significantly richer and they are more likely to have an internal migrant across all the three strata. Furthermore, in the Banjul and Kanifing and rural strata, the HFPS sample have a relatively larger household size compared to those without a valid mobile phone number. Finally, households in the HFPS are also more likely to have an international migrant in Banjul and Kanifing and the rural area strata than those without a mobile phone. The above findings suggest that ownership of valid mobile phone is significantly correlated with wealth.

Household members who were surveyed on the phone are slightly different from the face-to-face respondent[5]. On average, the phone respondent are mostly household heads (90 percent), spouse of the household head (5 percent), or child of the household head (4 percent). Furthermore, they are 45 years old on average; 25 percent are young (under 35 years old), mostly males (76 percent), and have some primary education (51 percent). Figure 1 compares the characteristics of the phone survey respondents with the characteristics of the heads of their households. Results suggests that household heads are statistically more likely to be males (82 vs 76 percent), older (48 vs 45 years), more likely to be without formal education (52 vs 0 percent) and less likely to achieve tertiary education (11 vs 30 percent). It is perhaps because of these differences that in some instances a representative of the household head is allowed to be interviewed instead of the household head.

Figure 1: Comparison between phone respondent and household heads
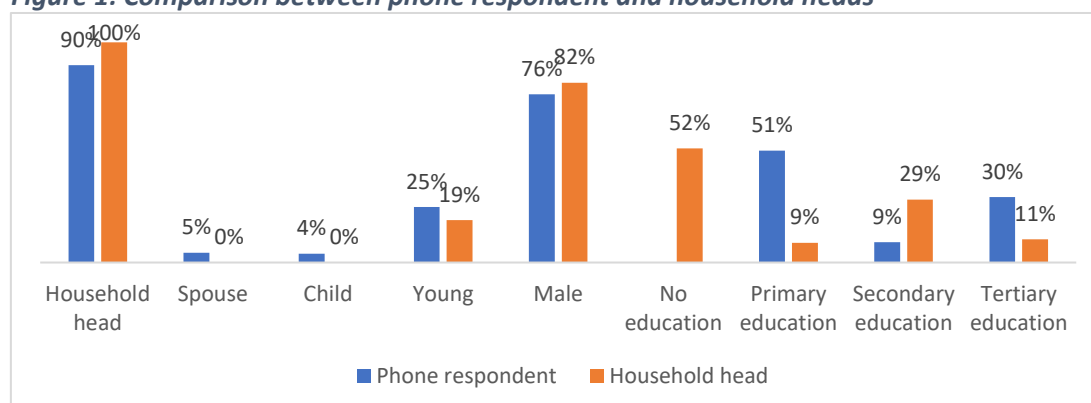
Source: own calculations HFPS data, 2020/21. All means are statistically different at 5 percent except for variable "male" which is significant at 10 percent.

4. Survey attrition

In this section, we describe the attrition rates in the survey. In the first wave (August 2020), about 1,437 households were successfully surveyed and in the latest wave (December 2021, 15 months after first wave), 886 households were surveyed. This suggests an overall attrition rate of 38 percent, averaging 16 percent per round of survey. Between waves 1 to 6, attrition was lower- under 15 percent- see Figure 2A. An increase of nearly 10 percentage point in attrition was reported in wave 7; and a further 8 percentage point increase in wave 10. The increase in attrition in wave 7 (July 2021) may in part be due to the commencement of the rainy season; whereas the increase in wave 10 (December 2021) may be due to the 2021 Presidential election events and a shorter data collection period. According to evidence elsewhere, similar attrition rates have been reported in similar surveys in other countries[6] (see Figure 2A). The main reasons

---

[5] The typical respondent in face-to-face surveys are the household heads.
[6] 42% attrition rate in Turkey after 18 months, 25% in Tanzania after 15 months and 15% attrition in Senegal (IPA, 2021) "Attrition in mobile phone panel surveys.- see
https://blogs.worldbank.org/impactevaluations/reducing-attrition-phone-surveys

for attrition in the HFPS include phone turned off (62 percent), nobody answering (29 percent) and refusal (5 percent). This suggests that connectivity problems are the main reason of attrition instead of refusal to participate in the survey.

Figure 2: Survey attrition and patterns across waves
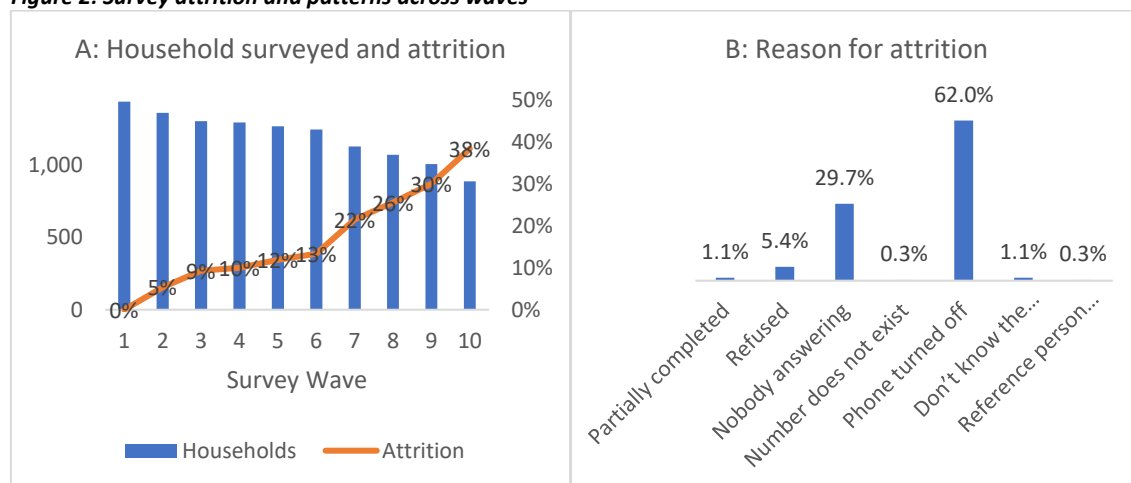
Attrition is correlated with gender, poverty status, education, and strata of residence of the respondent. Figure 3 provides differences between those who were not surveyed in at least one of the rounds (attrition) and those who participated in all the rounds. Among the strata group, attrition is highest in the other urban areas (61 percent), followed by the rural areas (31 percent) and Banjul/Kanifing (8 percent). The results suggest that males are less likely to miss at least one of the waves. Furthermore, those coming from the poorest 40 percent households and those in the rural areas are significantly more likely to participate in all rounds of the survey. In terms of education, respondents with secondary education have a higher attrition rates. Other variables including age, and relationship with the household head were not significantly different across attrition rates.

Figure 3: Differences between attrition and non-attrition respondents
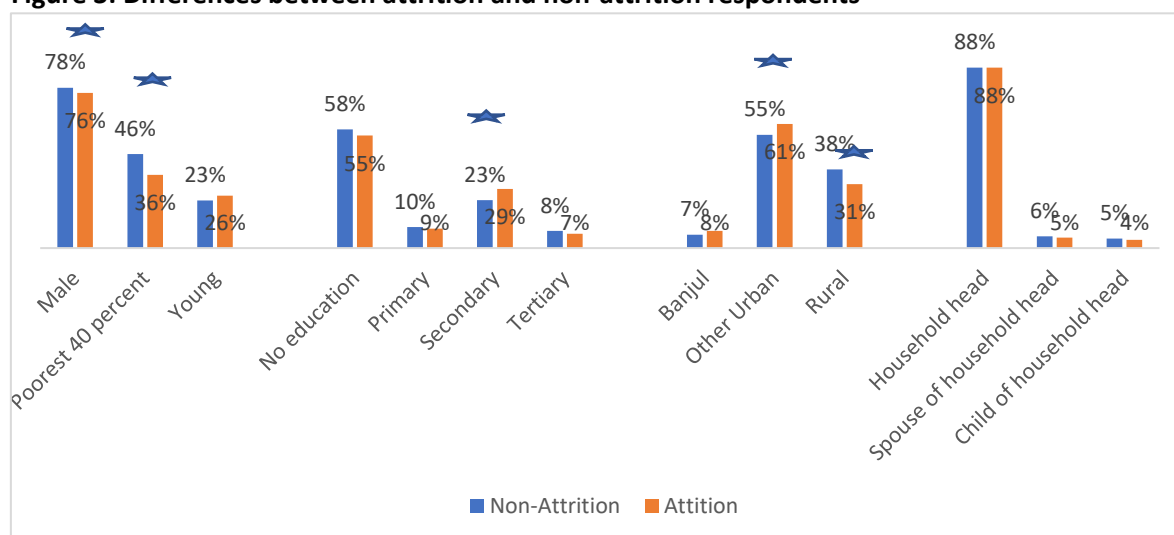
Source: own calculations HFPS data, 2020. ⭐ Indicate differences are statistically significant at 5 percent.

5. Conclusion and lessons learnt

The HFPS helped bridge critical data gaps at a time when traditional data collection techniques could not be conducted due to social distancing measures. It provided vital insights on the effect of the pandemic on various dimensions of household wellbeing -including labor market activities, income loss and food insecurity; during various phases of the pandemic. The survey also provided critical data to inform responses in health (especially on perception and uptake of vaccines) in education (learning in the new environment, needed reforms, impact of the pandemic on quality of learning, etc.) in social protection (access to social assistance, exposure to shocks and coping mechanisms); and in highlighting the effect of the pandemic on vulnerable population such as children and Persons With Disabilities (PWDs), among others.

Lessons learnt through the survey include:

1. Incentives provided to respondents may help minimize attrition but can introduce biases. Data on attrition in the survey shows that attrition rates among poorer households were lower than among richer households. This may in part be due to the incentives in the form of airtime which were offered to respondents. However, the effect of this attrition may be minimized by the fact that poorer households are less likely to own cell phones and hence lower attrition from them helps offset lower coverage of the poor in the sample.

2. Non-response can be minimized by asking respondents the best time to call them and assigning interviewers who speak the same language as them. In each round of the survey, respondents were asked about the best days and times to call. By leveraging this information, non-response was minimized. Additionally, information on language spoken by respondents was used to identify interviewers who spoke the same language. This helped minimize barriers in communication between interviewers and respondents.

3. By leveraging the flexibility in the design of the questionnaire for phone surveys, emerging data gaps can be identified and quickly filled. An important feature of phone surveys is the relative flexibility in the design of the questionnaire. Modules can be quickly added to monitor emerging events; and easily repeated to track trends overtime. Due to the nature of the COVID-19 pandemic, this feature of phone surveys was particularly helpful in generating critical data in a relatively short time to inform response to a rapidly evolving pandemic.

4. Although phone surveys are effective, there are limitations. Certain topics such as gender sensitive topics are difficult to collect through phone surveys- especially where one household member is interviewed. Additionally, only a limited number of questions can be included in any particular round of phone survey. Having a lengthy questionnaire introduces the risk of increasing the attrition rate. Despite these limitations, phone surveys can complement traditional data collection techniques- especially in times where such techniques cannot be utilized. Phone surveys are cheaper, more flexible and can provide data in a shorter time than typical face-to-face surveys.

5. In the future, better targeting of respondents (instead of relying on the household head or his/her representative) can be used to ensure a more gender balanced distribution of respondents. For instance, 75 percent of respondents interviewed in all the waves

were male. Possible alternatives to minimize these biases include leveraging the rich roster data to randomly identified household members to be interviewed instead of the household head or his/her representative.

Appendix A

Table A1: Summary Statistics: Statistical Differences between HFPS and Non-mobile phone sample

	(1) HFPS sample Mean	(2) No-mobile phone sample Mean	(3) T-test (1-2) p-value
Panel A: Banjul and Kanifing strata (household head/household characteristics)			
Age	45.862	44.031	0.108
Female	0.294	0.273	0.648
Household size	7.516	4.867	0.000***
Asset index	0.631	0.610	0.039*
Working-age population	0.891	0.908	0.320
Employment status	0.720	0.813	0.132
HH has external migrants	0.205	0.104	0.017*
Has attended school	0.993	1.000	0.083
Migrated internally	0.558	0.415	0.002**
Unemployed	0.042	0.004	0.027*
Observations	431	85	516
Panel B: Other urban areas strata (household head/household characteristics)			
Age	47.685	46.590	0.665
Female	0.146	0.216	0.871
Household size	9.552	8.810	0.822
Asset index	0.573	0.473	0.000***
Working-age population	0.885	0.865	0.220
Employment status	0.743	0.723	0.096
HH has external migrants	0.151	0.112	0.083
Has attended school	0.999	1.000	0.318
Migrated internally	0.454	0.360	0.023*
Unemployed	0.077	0.131	0.084
Observations	530	147	677
Panel C: Rural areas strata (household head/household characteristics)			
Age	49.103	51.551	0.561
Female	0.147	0.208	0.038*
Household size	11.030	10.550	0.001***
Asset index	0.484	0.423	0.000***
Working-age population	0.861	0.793	0.251
Employment status	0.527	0.495	0.015*
HH has external migrants	0.136	0.110	0.006**
Has attended school	1.000	0.997	0.318
Migrated internally	0.232	0.133	0.004**
Unemployed	0.261	0.281	0.093
Observations	509	223	732

Notes: HFPS sample represents households or individuals surveyed in the high frequency survey while no-mobile phone sample are those whose household do not have a valid listed phone in the LFS data are not part of this sample but surveyed in LFS 2018. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. There are 26 households from HFPS and 1 No-mobile phone household that are omitted due to missing asset index.
Source: HFPS 2020/21 and The Gambia Labor Force Survey, 2018