HIGH FREQUENCY SURVEY TO MONITOR COVID-19 IMPACT IN THE GAMBIA: TEN QUESTIONS TO UNDERSTAND THE SAMPLING PROCEDURES

<table>
<tr><th>QUESTIONS</th><th>ANSWER</th><th>COMMENTS</th></tr>
<tr><td>1. What is the sample frame?</td><td>2018 LFS</td><td>The 2018 LFS is the most recent household survey with phone contacts</td></tr>
<tr><td>2. What is the number of HH with valid phone number?</td><td>The 2018 LFS data set counts 5,987 household of which 5,531 household have phone numbers that are supposed to be correct. These numbers have 6 digits with 2,3,6,7 or 9 as first digit.</td><td>456 household have not phone numbers meaning that the corresponding variable is equal to 0 in the data set. These household are removed from the sample frame. Deep verification of the accuracy of the phone numbers is required to ensure that the sample is drawn from a cleaned sample frame.</td></tr>
<tr><td>3. What is the variable of interest?</td><td>The proportion of people employed in each household.<table><tr><th></th><th>Banjul&Kanifing</th><th>Other urban</th><th>Rural</th><th>Total</th></tr><tr><td>moyenne strate</td><td>26.0</td><td>22.1</td><td>16.4</td><td>20.3</td></tr><tr><td>Relative error</td><td>3.1%</td><td>3.6%</td><td>4.7%</td><td>2.2%</td></tr></table></td><td>We understand that the COVID affects the household through many channels. Due to the limitation of our sample frame (mostly oriented to the Labor market) and the needs to use the existing data as background information for the monthly phone surveys, the focus was made on the labor market indicator to identify the variable of interest. Since the LFS doesn’t collect accurately the income we propose alternatively the proportion of people employed in each household</td></tr>
<tr><td>4. What is the stratification level?</td><td>Three stratification levels are adopted: Banjul&Kanifing, other urban area and rural</td><td>Banjul is combined with Kanifing to have more observations. The sample frame is then distributed as following:<table><tr><th></th><th>Banjul&Kanifing</th><th>Other urban</th><th>Rural</th><th>Total</th></tr><tr><td>LFS sample with valid phone number</td><td>1,197</td><td>1,448</td><td>2,886</td><td>5,531</td></tr></table></td></tr>
<tr><td>5. What is the precision level to draw the sample?</td><td>7.5% of precision level</td><td></td></tr>
<tr><td>6. Distribution parameters?</td><td>T-student=1.96</td><td></td></tr>
<tr><td>7. Nonresponse rate?</td><td><table><tr><th></th><th>Banjul&Kanifing</th><th>Other urban</th><th>Rural</th><th>Total</th></tr><tr><td>LFS sample (theoric)</td><td>1,220</td><td>1,440</td><td>3,600</td><td>6,260</td></tr><tr><td>LFS sample with phone number</td><td>1,197</td><td>1,448</td><td>2,886</td><td>5,531</td></tr><tr><td>Response rate</td><td>1.9%</td><td>-0.6%</td><td>19.8%</td><td>11.6%</td></tr></table></td><td>The nonresponse rate including the phone availability is high a rural area. This presages a high sample size in rural area.</td></tr>
<tr><td>8. What is the sample size?</td><td>The minimal sample size is 1,524 based on the following parameters: nonresponse rate of 11.6%, low coverage of the variable of interest (20.3%) with a high standard deviation (17.1%).</td><td></td></tr>
<tr><td>9. Replacement rate?</td><td>A 25% replacement rate is considered to include the high attrition risk</td><td>To consider the high attrition risk, a replacement rate of 25% is considered corresponding to of 1,905hh. The idea is to call all the 1,905hh during the sensibilization stage. During the monthly interviews only 1,524hh will be covered and the additional 381hh will be used for</td></tr>
</table>

<table>
<tr><th>QUESTIONS</th><th>ANSWER</th><th>COMMENTS</th></tr>
<tr><td></td><td></td><td>replacement. Experienced based on recent surveys in Mali and Nigeria show a high attrition rate of 25%.</td></tr>
<tr><td>10. Which allocation for the sample by strata and to distribute the sample within the LGA?</td><td>Optimal allocation with the following results:
<table><tr><th></th><th>Banjul&Kanifing</th><th>Other urban</th><th>Rural</th><th>Total</th></tr><tr><td>Optimal allocation</td><td>414</td><td>544</td><td>566</td><td>1,524</td></tr><tr><td>Relative error</td><td>3.7%</td><td>3.0%</td><td>3.1%</td><td>1.9%</td></tr></table>Household within each LGA/Strate has the same chance to be drawn.</td><td>The code use is as following:
egen strateLGA=group(strate hh8)
sample StrateLGA_i , by(strateLGA)
count
ta strateLGA</td></tr>
</table>

PROGRAMME DE TIRAGE DES MENAGES DE L'ENQUETE COVID-19

***A defaut du revenu, nous avons retenu l'inverse du taux de dependance economique dans le menage comme variable d'interet

```
use "C:\Users\wb288526\OneDrive - WBG\Desktop\Djibril\COVID-19\Gambia\LFS_Final.dta" ,clear
sum nhhold_memb  //menage de 100 personnes???
preserve
collapse (first) hh5_2_1 hh5_2_2 weight count_of_unemployed employee_count self_employed_count nhhold_memb stratum hhold_addr, by( hh7 hh8 hh1 hh2)
count if hh5_2_1==. | hh5_2_2==.
ed
count if hh5_2_1==0 & hh5_2_2==0
save "C:\Users\wb288526\OneDrive - WBG\Desktop\Djibril\COVID-19\Gambia\sampleframe.dta",replace
restore
use "C:\Users\wb288526\OneDrive - WBG\Desktop\Djibril\COVID-19\Gambia\sampleframe.dta",replace
gen actif_occu=self_employed_count+employee_count
gen taux_depend=actif_occu*100/nhhold_memb
sum taux_depend
table hh8 [iw= weight], row col c(mean taux_depend)
tabstat taux_depend [aw=weight], by (hh8) stat (mean, count, p25, p50, p99, min, max )

count if hh5_2_1==0 & hh5_2_2==0
drop if hh5_2_1==0 & hh5_2_2==0
table hh8 [iw= weight], row col c(mean taux_depend)
tabstat taux_depend [aw=weight], by (hh8) stat (mean, count, p25, p50, p99, min, max )
clonevar phonevalide=hh5_2_1
replace phonevalide=hh5_2_2 if hh5_2_1==0
sum  phonevalide     //7 digits, premiers chiffres:2,3,6,7,9 ???


gen strate=.
replace strate=1 if hh8==1 | hh8==2
replace strate=2 if (hh8~=1 & hh8~=2) & hh7==1
replace strate=3 if hh7==2
ta strate, m

gen popweight= weight*nhhold_memb
svyset [pweight=popweight], strata(strate) psu(hh1)
svy: mean taux_depend, over(strate) // calcul de la moyenne et du standard error

estat effects, srssubpop // calcul deff et deft
estat sd, srssubpop  // calcul des ecarts-types

svy: mean taux_depend // calcul de la moyenne et du standard error

estat effects // calcul deff et deft
estat sd  // calcul des ecarts-types

**household sampling
egen strateLGA=group (hh8 strate )
lab def stratlag 1 "Banjul" 2 "Kanifing" 3 "Brikama urban" 4 "Brikama rural" 5 "MANSAKONKO urban" ///
6 "MANSAKONKO rural" 7 "KEREWAN urban" 8 "KEREWAN rural" 9 "KUNTAUR urban" 10 "KUNTAUR rural" ///
11 "JANJANBUREH urban" 12 "JANJANBUREH rural" 13 "BASSE urban" 14 "BASSE rural"
lab val strateLGA stratlag


gen EchantStrat= 414 if strate==1
replace EchantStrat=544 if strate==2
```

```
replace EchantStrat=566 if strate==3
gen test=1
bysort strate: egen MenStrat=count(test)
ta MenStrat
gen SampStratWeight=EchantStrat/MenStrat
ta SampStratWeight

bysort strateLGA : egen PopStratum= total (weight*nhhold_memb)
bysort strate : egen PopStrate= total (weight*nhhold_memb)
gen SampleSratLgaWeight =PopStratum/PopStrate

gen SampDetLga=SampStratWeight*SampleSratLgaWeight
ta strateLGA SampDetLga


preserve
keep if strate==1
sample 34.6047589, by (strateLGA)
keep hh7 hh8 hh1 hh2 weight stratum hhold_addr taux_depend phonevalide EchantStrat MenStrat SampleSratLgaWeight strateLGA
save "C:\Users\wb288526\OneDrive - WBG\Desktop\Djibril\COVID-19\Gambia\sampleStrate1.dta",replace
restore


preserve
keep if strate==2
sample 37.5690413, by (strateLGA)
keep hh7 hh8 hh1 hh2 weight stratum hhold_addr taux_depend phonevalide EchantStrat MenStrat SampleSratLgaWeight strateLGA
save "C:\Users\wb288526\OneDrive - WBG\Desktop\Djibril\COVID-19\Gambia\sampleStrate2.dta",replace
restore



preserve
keep if strate==3
sample 19.608552, by (strateLGA)
keep hh7 hh8 hh1 hh2 weight stratum hhold_addr taux_depend phonevalide EchantStrat MenStrat SampleSratLgaWeight strateLGA
save "C:\Users\wb288526\OneDrive - WBG\Desktop\Djibril\COVID-19\Gambia\sampleStrate3.dta",replace
restore

use "C:\Users\wb288526\OneDrive - WBG\Desktop\Djibril\COVID-19\Gambia\sampleStrate1.dta",clear
append using "C:\Users\wb288526\OneDrive - WBG\Desktop\Djibril\COVID-19\Gambia\sampleStrate2.dta"
append using "C:\Users\wb288526\OneDrive - WBG\Desktop\Djibril\COVID-19\Gambia\sampleStrate3.dta"
gsort strateLGA
save "C:\Users\wb288526\OneDrive - WBG\Desktop\Djibril\COVID-19\Gambia\SampleFINAL.dta",replace
stop
```