Sampling Methodology for Malaysia's Phone-Based High-Frequency (HiFy) Monitoring of COVID-19 Impacts on Households

I. SAMPLING FRAME

The Malaysia HiFy survey is a computer-assisted telephone interviewing (CATI) based survey, with a mobile-only frame generated via random digit dialing (RDD). Documentation from the Malaysian Communications and Multimedia Commission (MCMC) was used to identify all active mobile provider codes. All possible subscriber combinations were generated in D Force Sampling's Reactive User Interface Database (DRUID), which houses the complete sampling frame. From this database, complete random telephone numbers were sampled. There was a total of 102,780,000 possible mobile numbers in the sampling frame, which were not stratified. [1]

II. SAMPLING METHOD

Target Population: Telephone accessible population of Malaysian citizens, aged 18+

Target Sample Size: 2,200 Malaysians aged 18+ who either make financial decisions, contribute or are knowledgeable on household finances.

The total number of phone numbers sampled for Round 1 was 33,894. Of these, 2,210 interviews were completed. In Round 2, all of 2,210 respondents who were successfully interviewed in Round 1 were followed up. Of these, 1,047 interviews were completed.

The sampling methodology for this study was as follows:

1. **Simple random sample of numbers**: In Round 1, mobile numbers are generated via simple random sampling and are sampled through DRUID.
   a. All numbers are sampled without replacement within the survey wave.
   b. However, numbers can be reintroduced into the sample frame for subsequent waves, should there be continued studies.

---

[1] There are more than 18 mobile providers in the sampling frame.

2. **Preliminary auto-dialer filtering:** In Round 1, an auto-dialer filtering procedure was implemented; once the sample is drawn in the form of replicates (subsamples) of n=10,000, the telephone numbers are filtered using an auto-dialer to determine each number's working status.

   a. Up to two automated call attempts per number are done to determine whether the number is active or not, recording a disposition for each attempt. A minimum of two hours between each attempt is specified in the system.

   b. All numbers that yield a working call disposition for at least one of the two filtering attempts are then passed to the CATI center human interviewing team.

   Round 2 did not undergo auto-dialer filtering as the sample comprised of participated respondents in Round 1.

3. **Method of respondent selection**: The devices were assumed to be personal, and therefore the person who answered was the selected respondent.

   a. In Round 1, screener questions asking the age of the person who answers were used to ensure that the respondent is at least 18 years old, as well as within the capacity of either contributing, making or have knowledge on household finances.

   b. In Round 2, screener questions were used to ensure that the respondent of the recorded phone number in Round 1 was the same person who participated in the Round 1 of the survey.

4. **Respondent and interviewer matching**: Each selected respondent was offered the choice to complete the survey with an interviewer of matching ethnic background and/or language capability (Malay, Chinese, or English-speaking interviewers).

5. **Incentives**: Respondents were offered incentive tokens with a small value in the form of either a prepaid telephone reload, or postpaid telephone reload.

6. **Callbacks:**

   a. In Round 1, up to five callbacks were attempted for each filtered telephone number to obtain a completed interview. Therefore, the maximum number of permitted contact attempts totals 7 (two autodialed filtering attempts

and five attempts to establish contact with an eligible respondent and for completing the survey).

    b. In Round 2, up to eight callbacks were attempted for each telephone number to obtain completed interviews. Callback intervals for both rounds are half an hour for busy lines and every three+ hours for answering machines / not reached.

7. **Respondent substitution:** If the interviewers fail to complete an interview with the selected respondent after all callbacks are exhausted, they are not allowed to perform an in-house substitution. Each call attempt disposition will be properly entered, and interviewers will move onto the next randomly generated telephone number.

8. **Replicate structure and release protocols:**

    a. Phone numbers are sampled and delivered in subsamples called replicates.

    b. The standard of effort outlined in the Callbacks step above is made per replicate before the next is released. This procedure ensures that the same level of effort is made for all phone numbers, diminishing the bias against households and respondents that are less likely to answer the phone on the first attempt.

9. **Weighting:** No quotas are implemented during field; rather, the achieved sample is weighted post-field with adjustments including a base weight, non-response weight, raking & trimming, and rescaling, explained in full detail in the following section.

III. WEIGHTING

A full probability-based sampling methodology was implemented. In Round 1, the complete weight is for the entire sample adjusted to 2019 population estimates from annual intercensal population projections from the Department of Statistics Malaysia (DOSM) on the designated parameters, including gender, age, state, and ethnic group. The estimates excluded the portion of the population that is not part of the target universe, particularly for age group and citizenship status. Assuming a simple random sample, with $p=0.5$ and $n=2,210$ at the 95% CI level yields a margin of sampling error

(MOE) of 2.09 percentage points. Incorporating the design effect into this estimate yields a margin of sampling error of 2.65% percentage points.

In Round 2, the complete weight is for the entire sample adjusted to the 2021 population estimates from DOSM's annual intercensal population projections. Assuming a simple random sample with p=0.5 and n=1,047 at the 95% CI level yields a margin of sampling error (MOE) of 3.803 percentage points. Incorporating the design effect into this estimate yields a margin of sampling error of 3.54 percentage points.

The weighting scheme for Round 1 was developed with the following adjustments via weighting:

1. **Base Weight:** a base weight was calculated as the inverse of the probability of a number being dialed.
2. **Non-response Weight:** a non-response weighting adjustment was performed using a weighting-class adjustment by the inverse of Response Rate 3 (as defined by AAPOR) by the sample design stratum. Phone numbers that were removed via pulsing are considered part of Not Eligible for this weighting adjustment.
3. **Raking and Trimming:** a post-stratification weighting adjustment was performed using the aforementioned benchmark source for national population figures. An outlier analysis of the weights was then performed, with those being beyond 3 standard deviations of the mean being trimmed. This was done as an iterative process (raking, trimming, and raking again) until weights were stable within a comfortable max weight.
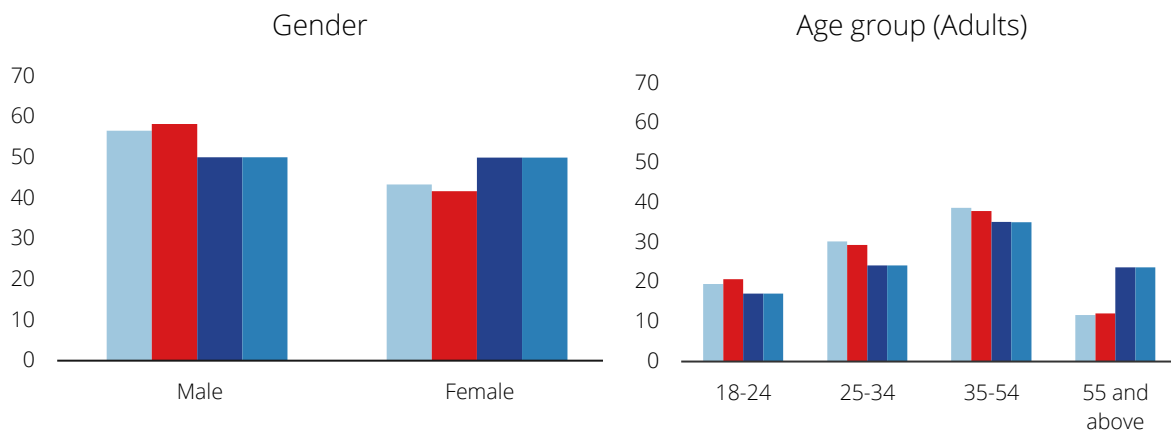4. **Rescaled:** weights are also delivered in a rescaled format.

The weighting scheme for Round 2 was developed using adjustments in the form of raking and trimming, as well as weights rescaling.
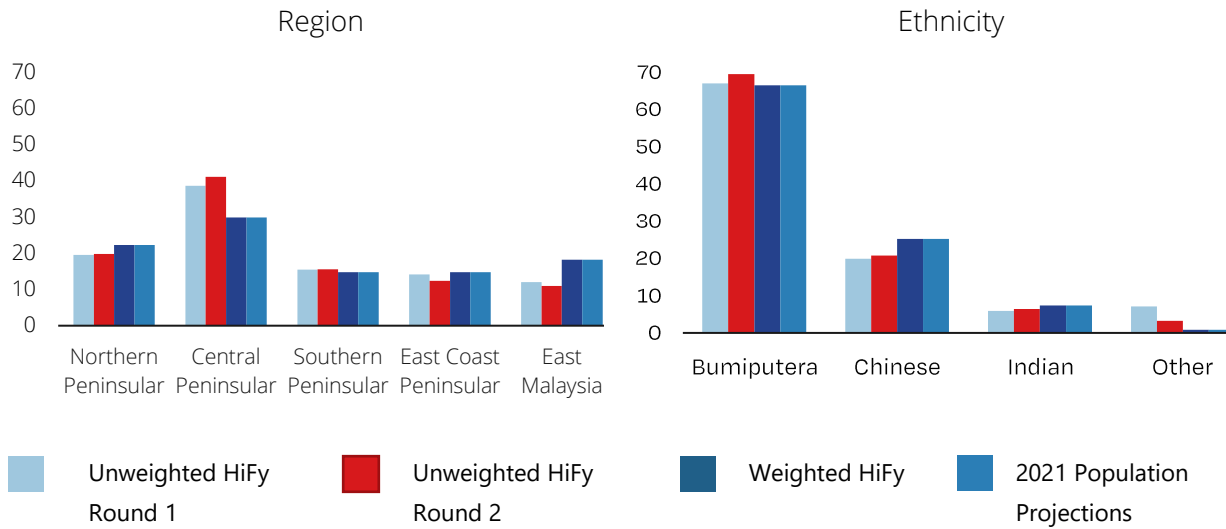
IV ATTRITION

The attrition occurred when respondents from Round 1 were not able to be re-interviewed in Round 2. The attrition rate for the HiFy survey in Round 2 was 52.6%, with 1,047 respondents who had previously participated in Round 1 were successfully re-interviewed in Round 2. Most of the dropouts in Round 2 was due to unanswered calls (62% of all dropouts) or because their phones were uncontactable (14% of all dropouts). 14% of the

dropouts refused to participate in the second round. A test to investigate whether attrition was random showed that the dropped respondents were not associated with gender, age group, household income groups, region, or ethnicity. However, there was a weak association between survey participation with respondents' education, as respondents having post-secondary education were more likely to participate in the Round 2 than those with lower education. While such education bias couldn't be taken into account in the weight calculation due to limited data availability in the DOSM's annual inter-censal population projections, the distribution of respondents by education levels remains comparable with the population estimates of head of households' education from Household Income Survey (given the screening questions for respondents should be involved in household financial decisions, it is reasonable to expect that, *ceteris paribus*, those with more education are more likely to be involved in such decisions and hence participate in the survey). Further, weights are also implemented to ensure that the dropouts bear no significant impact on the representativeness of our panel samples, along key demographics (Figure 1).

Figure 1 Key demographics in different HiFy survey rounds, unweighted and weighted versions, compared with DOSM's 2021 population projections

Region | **Ethnicity**

Legend:
- Unweighted HiFy Round 1
- Unweighted HiFy Round 2
- Weighted HiFy
- 2021 Population Projections

V. ADDITIONAL NOTES

Sampling Parameter:

Prior to using the HiFy Round 1 data for in-depth analysis, the following sampling set-up can be used in the Stata software:

```
gen stratif = (state * 10000) + (sex * 100) + (age_group * 10) +
ethnic_group

svyset s0_m_01 [pw=finalwgt2], strata(stratif) vce(linearized)
singleunit(certainty) clear
```

where,

- *finalwgt2* is the overall weight
- *stratif* is the combination of several sampling units used, including state, gender, age group, and ethnic group

For the HiFy Round 2 data, one can implement the following sampling set-up prior to data analysis:

```
gen stratif_w2 = (state * 10000) + (sex * 100) + (age_group * 10) +
ethnic_group
```

```
svyset s0_m_01 [pw=finalwgt2], strata(stratif_w2) vce(linearized)
singleunit(certainty) clear
```

where,

- *finalwgt2* is the overall weight
- *stratif_w2* is the combination of several sampling units used, including state, gender, age group, and ethnic group