

Technical Documentation: Lebanon Pulse Survey

Overview

The Lebanon Pulse Survey (LPS) responds to a persistent lack of timely, consistently collected micro-level data on household economic conditions in Lebanon. In contexts marked by economic volatility, limited institutional data, and rapidly changing living conditions, aggregate indicators alone do not fully capture how households experience and respond to economic change. The survey collects objective data on household socioeconomic conditions alongside complementary information on perceptions, expectations, and concerns about the economy.

Subjective indicators provide insights into households' well-being, future economic and social decisions, and the risks households themselves see as most pressing. When collected regularly, they offer a continuous dataset that helps track changes in household welfare, economic sentiment, exposure to shocks, and expectations for the future.

The survey follows the logic of other data initiatives that use household- or individual-level surveys to track social and economic conditions over time. These include rapid-response data collection projects used during periods of crisis or disruption, such as the World Bank's high-frequency Household Monitoring Systems, which were deployed in multiple countries to monitor household welfare during the COVID-19 pandemic. Other initiatives are of a more structural nature, such as the U.S. Census Bureau's Household Pulse Survey and the University of Michigan Surveys of Consumers, which were developed to rapidly measure emerging social and economic conditions affecting households. In the region, projects such as the Arab Barometer demonstrate the value of repeated survey-based measurement for understanding public attitudes and socioeconomic conditions across countries.

By producing regularly updated, comparable, and analysis-ready microdata, the LPS aims to support empirical research, policy analysis, and public-interest reporting on the evolving status of the Lebanese economy and its effects on households. Its repeated design allows researchers to track changes across survey rounds while maintaining a standardized methodological framework for sampling, data collection, and quality standards.

Section 1. Survey Design and Sampling Methodology

This section outlines the design and implementation of the LPS. It describes the target population, sampling approach, and questionnaire structure used to collect data that is intended to be representative of households across Lebanon. The methodology is designed to ensure adequate geographic and socio-economic coverage, while maintaining consistency across survey rounds to allow comparability over time.

Target Population and Eligibility Criteria

The target population consists of adult residents of Lebanon. Residency is defined as having resided in Lebanon during the previous 12 months, irrespective of nationality or legal residency status. Participation is restricted to adults who are aged 21 years or older. The sample excludes non-residents and short-term visitors such as tourists, individuals on short-term work trips, people on medical stays, and members of diplomatic missions. Institutionalized and incarcerated individuals are not sampled in the study.

Survey Periodicity and Structure

The survey follows a biannual repeated-cross-section design with a short-panel component. A fresh sample is randomly drawn for the main survey round conducted at the beginning of each year. In July, a randomly selected subsample of respondents from the annual main-round sample is re-contacted and re-interviewed as part of the follow up round. Respondents interviewed in both rounds can be linked across survey waves using a unique individual survey identifier.

Although the sampling design could allow respondents to participate in multiple annual survey cycles, respondents are linked only within the same survey year. Linkages are maintained only between the main round and the July follow-up round within the same survey year.

Sampling Frame, Randomization and Representativeness

The Republic of Lebanon does not maintain an active national census. In the absence of an accurate population count, the baseline sampling framework relies on recent national and sub-national population estimates. This includes population estimates published in the yearly statistical bulletin of the Ministry of Public Health, and estimates produced by the Central Administration of Statistics. The latest population estimates are included in Appendix A.

A fresh random sample is selected at the beginning of every calendar year. Sample recruitment strategy relies on a Random Digit Dialling (RDD) procedure. Mobile phone numbers are randomly generated within valid Lebanese mobile numbering ranges, and eligible respondents are interviewed by trained data collectors via WhatsApp calls.

A list of 100,000 unique Lebanese mobile phone numbers with valid prefixes¹ was randomly generated. This ensures coverage across telecom providers and both older and newer prefix ranges. Given high mobile phone penetration in Lebanon, with most adults owning a mobile phone (PEW Research Center, 2019) and near-universal network coverage (ITU, 2023), this approach provides broad population reach.

Selected phone numbers are then screened using WhatsApp to identify active and valid numbers. WhatsApp is widely used across Lebanese and non-Lebanese populations in Lebanon and serves as a major communication platform in the country, making it a practical tool for identifying active mobile numbers associated with WhatsApp accounts.² To reduce sample selection based on economic activity, data collection activity is conducted over an extended daily schedule during and after working hours, and during regular working weekdays as well as weekends.

Sample Size, Survey Precision, and Power Calculation

The target sample size was determined based on a 95% confidence level and a margin of error of 2.5 percentage points for national-level estimates³. The latest population estimate from the Ministry of Public Health (Appendix A) places the resident population aged 20 years or older in Lebanon at around a total of 4,111,906 residents. Given this population size, the sample size was calculated using the formula for large or “effectively infinite” populations, as the finite population correction would have a negligible effect on the final estimate.

The required sample size under simple random sampling is calculated as follows:

$$n = \frac{Z^2 \times p(1 - p)}{E^2}$$

¹ For the purpose of constructing the mobile number sampling frame, Lebanese mobile phone numbers consist of a two-digit prefix followed by six digits. The eligible mobile prefixes are 03, 70, 71, 76, 78, 79, 81, 82, 83, 84, 85, 86, 87, 88, and 89.

² UNDP (2024) describes WhatsApp as a pivotal communication platform in Lebanon. The report cites Pew Research Center (2019), which found that WhatsApp was the preferred messaging platform among Lebanese adults, with a usage rate of 84 percent, and UNDP (2020), which reported that 78 percent of Syrian households in Lebanon used WhatsApp for daily communication.

³ A pilot conducted in February 2026 included 182 respondents with an initial screening of ~3000 mobile numbers for WhatsApp availability. Of these, 66% were either not registered on WhatsApp or could not be verified as active numbers. Among the remaining contacts with confirmed WhatsApp access, 18% resulted in successful interviews, while 38% declined participation and 44% did not respond.

Where n is the required sample size, Z is the Z-score for the chosen confidence level, p is the assumed population proportion, and E is the margin of error.

For a 95% confidence level, $Z=1.96$. In the absence of prior estimates for the study variables (e.g., economic sentiment, migration intentions, and related attitudes), **p is set to 0.5**, as this maximizes variance in a binomial distribution and yields the most conservative sample size estimate. This is standard practice when population proportions are unknown (Cochran, 1977). The required sample size under simple random sampling is therefore 1,537.

Since the survey is designed to support analysis across demographic groups and regions, and because the realized survey design may involve stratification, weighting, and differential response patterns, the sample size is inflated by an expected design effect (DEFF)⁴. An ex-ante design effect of 1.5 is used as a conservative planning assumption to account for the expected loss of precision resulting from survey design and implementation. The adjusted completed sample size is 2,305.

The survey therefore targets a completed sample of approximately 2,300 to 2,500 respondents per round. This target balances statistical precision, the need for subgroup analysis, and the practical cost of repeated data collection. It provides relatively strong precision for national-level estimates and for major demographic groups, while estimates for smaller regional or demographic cells will be interpreted with wider confidence intervals.

The target sample size provides sufficient statistical power for the main comparative analyses planned for the survey. Given the mix of binary outcomes and Likert-style perception and expectation measures, the survey is powered to detect substantively meaningful differences across large demographic subgroups and across survey rounds. Statistical power will be strongest for full-sample comparisons and for large demographic groups, such as comparisons by gender, broad geographic area, and citizenship status. Comparisons involving smaller regional or demographic cells will have lower precision and will require larger differences to be statistically distinguishable.

The study team monitors field progress throughout data collection to ensure sufficient completed interviews across key demographic and regional groups. Post-survey checks are also conducted to assess representativeness of the achieved sample relative to population benchmarks.

⁴ The design effect captures the expected loss of statistical precision relative to a simple random sample of the same size.

Survey Instrument and Questionnaire Structure

The survey uses a modular questionnaire developed by the Lab research team, to measure objective household conditions and subjective economic perceptions. The questionnaire structure differs by survey round: an extended instrument is administered during the main annual round, while a shorter instrument is used for the July follow-up survey.

The survey instruments consist primarily of closed-ended questions with standardized response categories, facilitating consistent administration across respondents and comparability across survey rounds. The questionnaires also include selected open-ended questions that allow respondents to provide unstructured information on their opinions and concerns. The instruments are organized into four core modules:

1. Demographic and socio-economic characteristics
2. Recent economic experiences (income, employment, shocks)
3. Future expectations (income, employment, economy)
4. Open-ended qualitative questions

Likert-scale questions are included to measure perceptions and economic sentiment in a structured and comparable way. The scale contains five possible options including a neutral one when respondents feel neither positive nor negative about a topic. Respondents can opt out of responding to most of the questions by indicating that a question is not applicable to their experience.

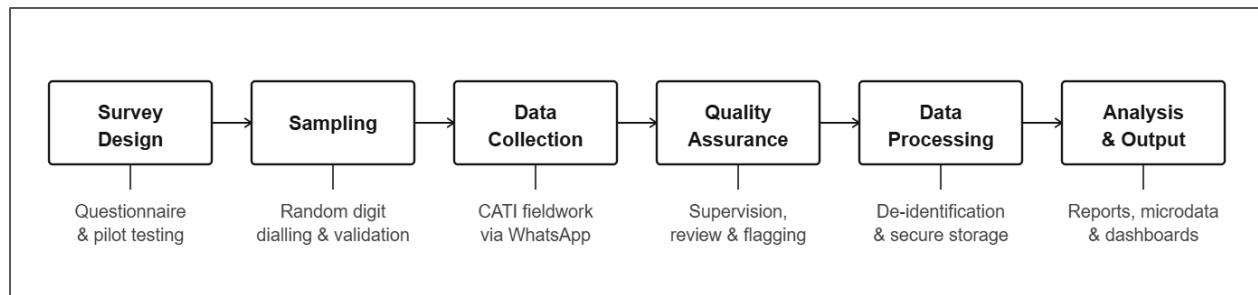
The survey design includes recall periods of 6 months or one year that are explicitly defined within each module to reduce variation in interpretation across respondents.

Key questions are intentionally repeated or rephrased at different points in the questionnaire to improve internal consistency and reduce the impact of response bias and order effects. The sequencing of questions follows a structured flow, starting with neutral demographic questions and gradually moving toward more sensitive economic topics to reduce respondent fatigue and improve response quality.

The questions included in the core modules were tested during the pilot survey conducted in calendar year 2026. Based on the pilot results, the questionnaire was revised to improve clarity, consistency, respondent comprehension, and alignment with the survey's measurement objectives.

The modular structure of the survey instrument also allows for the inclusion of supplemental sections in specific survey rounds. These sections may be added to capture emerging issues, context-specific development, or topics of immediate policy and research relevance at the time of survey administration.

Figure 1. *End-to-end survey design and data pipeline*



Section 2. Ethical Considerations and Data Security

The survey is implemented in accordance with standard ethical principles for social science research, including voluntary participation, informed consent, confidentiality, and protection of respondent privacy. Data collection and management protocols are designed to minimize risks to participants and to ensure that identifiable information is handled securely and accessed only by authorized members of the research team. The study protocol was reviewed and approved by the Institutional Review Board at the American University of Beirut under protocol ID SBS-2025-0567.

Informed Consent

Participation in the LPS is voluntary and based on informed consent. Prior to the start of each interview, respondents are informed about the purpose of the study, the expected duration of the interview, the types of information to be collected, and the intended use of the data. After verbal consent is obtained, respondents are sent a written informed consent statement, provided in both Arabic and English. Respondents may decline to answer any question or terminate the interview at any point without penalty.

Confidentiality and Privacy

All information collected through the LPS is treated as confidential. Respondent identities are protected throughout the data collection, processing, analysis, and reporting stages. Personally identifiable information, including phone numbers and any contact information

required for follow-up, is stored separately from survey responses and is used only for survey administration and respondent re-contact where applicable.

Each respondent is assigned a unique randomly generated survey identifier used to link records within the same survey year, when respondents participate in both the main and follow-up rounds. Phone numbers and other direct contact information are retained only for the duration necessary to complete the within-year follow-up survey and are discarded at the end of each survey year after completion of the follow-up round. All data collectors and research team members involved in the research receive training on ethical data use, respondent confidentiality, privacy protection, and secure handling of survey data. The training emphasizes the responsible use of personally identifiable information, restrictions on unauthorized data sharing, and procedures for protecting respondent anonymity throughout the research process.

Data Security and Access Control

Survey data are collected using a secure web-based data collection platform. After completion of data collection, survey data are downloaded and stored in a dedicated shared folder on the University's approved storage system. The cloud storage system used for the project is the institutional platform approved by AUB Information Technology services and designated for secure data storage and collaboration. Access permissions are managed by the research team and limited to personnel with legitimate research or operational need to access the data.

Responsible Data Use

The survey data are used exclusively for academic research, policy analysis, and public-interest reporting related to the objectives of the LPS. Access is restricted to authorized Lab personnel and collaborators, and use of the data is governed by the confidentiality and data security procedures described above.

All analyses are conducted using de-identified datasets. Direct identifiers are not used in statistical analysis, reporting, and dissemination outputs. All outputs derived from the survey are presented in aggregate form only and are designed to prevent identification of individual respondents.

Any external data sharing, where applicable, is subject to prior review, appropriate de-identification, and compliance with the study's ethical approval, institutional data governance requirements, and applicable privacy standards.

Section 3. Data Collection Procedures

Data collection is conducted using standardized field protocols designed to ensure consistency across respondents, data collectors, and survey rounds. The procedures cover respondent contact, informed consent, questionnaire administration, data entry, field supervision, and quality assurance. These protocols are designed to minimize measurement error, reduce interviewer effects, and maintain consistency across survey rounds.

The survey is administered using a Computer-Assisted Telephone Interviewing (CATI) approach conducted via WhatsApp, with KoboToolbox serving as digital data collection platform. Interviewers administer the questionnaire through KoboToolbox, which incorporates standardized interview scripts, programmed skip logic, validation rules, and consistency checks to support uniform data collection.

Prior to fieldwork, interviewers complete a four-week training with the data Lab coordinator and research lead. Training covered the survey objectives, questionnaire content, respondent contact protocols, informed consent procedures, ethical data collection practices, CATI procedures, and standardized data-entry protocols.

Interviews are conducted from the Lab premises, with at least one supervisor present during fieldwork. Supervisors monitor field work progress, including completion rates, interview duration, and overall data quality. They also review flagged observations and follow up with interviewers when inconsistencies, missing information, or deviations from field protocols are identified.

Interview Protocol

Telephone numbers are selected using the random digit dialling procedure described in Section 1. Randomly generated telephone numbers are assigned to interviewers using standardized fieldwork procedures designed to distribute workload evenly and monitor call outcomes.

Data collectors contact the assigned numbers following the standardized CATI protocol. At the start of each call, the data collector introduces the survey, verifies that the respondent is eligible to participate, and administers the informed consent script. If the respondent provides consent, they are asked whether they prefer to complete the interview immediately or schedule it for a later time.

For each completed or attempted interview, the data collector records the relevant call outcome using standardized response codes. These include completed interview, partial interviews, refusals, ineligible respondents, non-response, and scheduled callbacks. Response outcomes are reviewed periodically during fieldwork and retained for use in monitoring survey progress, assessing nonresponse patterns, and constructing survey weights when needed.

For respondents who consent to participate, KoboToolbox generates a unique random individual identifier. The data collector records this identifier in a restricted tracking file together with the respondent's phone number. This tracking file is used only for survey administration, including callback scheduling and within-year follow-up interviews. It is stored separately from the analytical dataset and is discarded at the end of the year in accordance with the confidentiality and data security protocol described above.

Interviews are conducted primarily in Arabic. Interviewers may switch to English or French when preferred by the respondent. This language flexibility is used to improve respondent comprehension, reduce miscommunication, increase representativeness of residents from diverse linguistic backgrounds, and support standardized administration across respondents with different language preferences.

Call Attempts and Nonresponse Protocols

Each sampled mobile number is called multiple times at different times of day and on different days of the week to increase the likelihood of reaching eligible respondents and to reduce contact-related nonresponse. If a respondent cannot be reached on the initial attempt, additional call attempts are scheduled according to a predefined callback protocol.

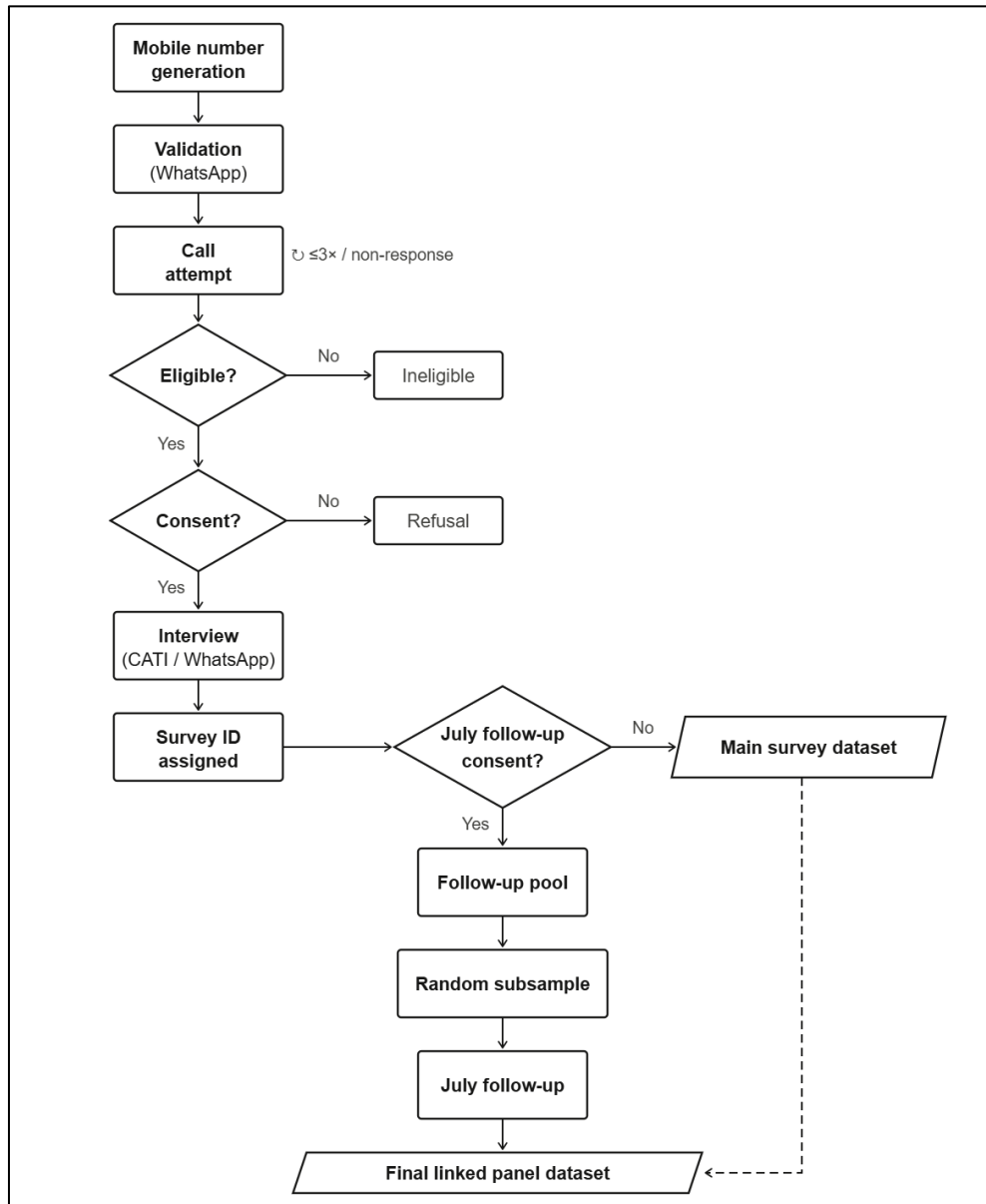
Numbers that remain unreachable after three call attempts are classified as nonrespondents. Respondents who decline participation are classified as refusals and are not re-contacted unless they explicitly request to be contacted at a later time. Partial interviews are reviewed by the research team to determine whether they meet the minimum completeness criteria for inclusion in the final dataset.

Fieldwork Monitoring and Quality Assurance

Fieldwork is monitored continuously by senior members of the research team to ensure consistent implementation of the survey protocol and high data quality. During data collection, at least one senior team member is present with the data collectors to provide real-time supervision and operational support. Supervisors monitor interviewer adherence to

standardized interview procedures, interview completion rates, interview duration, refusal rates, item nonresponse, and the overall quality and consistency of submitted records.

Figure 2. *Field operations flow*



Field supervision also includes regular communication between interviewers and the research team. Interviewers are encouraged to report questions that respondents find difficult to understand, response categories requiring clarification, and any recurring issues encountered during data collection. Where appropriate, this feedback is translated into standardized guidance that is communicated to all interviewers to promote consistent administration of the questionnaire across the field team.

Quality assurance is supported by both automated and manual review procedures. The electronic questionnaire incorporates built-in validation rules, programmed skip patterns, range restrictions, and consistency checks to reduce data-entry errors and minimize invalid responses at the point of collection.

Submitted records are reviewed periodically by the research team to identify missing values, inconsistent responses, unusually short interview durations, high item nonresponse, or other irregular patterns. Records flagged during these reviews are examined by the supervisory team and, where necessary, discussed with the relevant interviewer to determine whether clarification, correction, or exclusion from the final analytical dataset is warranted.

Section 4. Data Processing and Weighting

Raw survey data are processed through a standardized workflow designed to produce an analysis-ready dataset. The workflow includes data export, systematic data cleaning and validation to identify and correct inconsistencies, and the generation of survey weights. These procedures are documented for each survey round to ensure reproducibility and consistency across survey rounds.

Data Export and Secure Storage

Upon completion of fieldwork, raw survey data are exported and stored in a dedicated project folder on a cloud-based storage system. The raw data export is retained as an unmodified source file to preserve the original fieldwork record. Subsequent cleaning, recording, variable construction, and analysis are conducted on separate working copies to ensure that all data-processing steps are reproducible and can be traced back to the original raw dataset.

Project files are organized using a standardized folder structure that separates raw data, cleaned data, analysis-ready datasets, code files, documentation, and output files. Where applicable, file names include the survey round, export date, and version number to support version control and reduce the risk of accidental overwriting or confusion across survey waves.

As an additional data protection measure, a backup copy of the survey data is stored on an encrypted physical hard drive. The hard drive is kept in a locked office cabinet and is accessible only to authorized members of the Lab. These procedures are designed to protect the data against unauthorized access, accidental loss, unintended modification, and improper disclosure.

Data Cleaning and Validation

Data processing begins with systematic cleaning and validation of records. Given the interviewer-administered mode of data collection, initial validation checks focus on potential data-entry errors, and deviations from programmed skip patterns or expected response ranges.

Logical consistency checks are applied to identify responses that are internally contradictory or unlikely given the structure of the questionnaire. For example, observations are flagged when a respondent reports zero working hours while being classified as employed full-time. Flagged observations are reviewed by the research team to determine whether they reflect data-entry errors, interviewer misclassification, respondent misunderstanding, or valid but unusual cases.

Where corrections are warranted and can be verified from available information, the cleaned dataset is updated accordingly. Cases that cannot be resolved with sufficient confidence are retained with appropriate missing-value or flag indicators, rather than being modified without verification.

Treatment of Partial and Incomplete Interviews

An interview is classified as complete when the respondent reaches the end of the questionnaire and provides valid responses to the core modules required for analysis. Interviews that are terminated before completion are classified as partial interviews.

Partial interviews are assessed based on the point at which the interview was discontinued, the extent of missing information, and the availability of responses to key demographic and socioeconomic variables. Partial interviews that contain sufficient information for analysis are retained in the dataset. Records with insufficient information, substantial missingness in core variables, or unresolved inconsistencies are excluded from the analysis-ready dataset.

Variable Construction and Coding

After the completion of data cleaning and validation, raw survey responses are transformed into standardized analysis-ready variables. This process includes recoding response categories, variable labelling, constructing derived indicators, and applying consistent missing-value coding.

Categorical variables are coded using standardized numeric codes with corresponding value labels. Demographic variables such as gender, marital status, religion, governorate of residence, and nationality are recoded into harmonized categories that remain consistent across survey rounds. Open-ended responses are retained in their original language and reviewed separately from structural variables.

Missing values are coded using a standardized missing-value protocol. Distinct codes are used to distinguish between valid nonresponse, refusal, “do not know” responses, not-applicable items resulting from skip patterns, and missing values due to incomplete interviews or data-entry issues.

All constructed variables, recoding decisions, and derived indicators are documented in a data dictionary accompanying each survey round. The data dictionary includes variable names, labels, response categories, coding rules, and missing-value definitions.

Weight Construction

Survey weights are constructed to improve the representativeness of survey estimates and to account for the sampling and fieldwork design. The weighting procedure accounts for the random digit dialling sampling process and adjusts for differential contact and response probabilities.

The initial base weight reflects the probability of selection from the generated mobile-number sampling frame. In the final stage, weights are calibrated to external population benchmarks to reduce discrepancies between the achieved sample composition and the target population distribution across key demographic and geographic characteristics.

Population benchmarks are drawn from official population estimates published by the Lebanese Ministry of Public Health in its annual statistical reports and population estimates produced by the Central Administration of Statistics. These sources are used to define calibration margins for variables such as sex, age group, governorate of residence, nationality, or other relevant characteristics.

The distribution of final weights is reviewed to identify extreme values and assess the effective sample size. Any trimming, normalization, or additional calibration decisions are documented in the survey-round report.

Final Analysis Dataset

The final analysis dataset is produced after the completion of data cleaning, validation, variable construction, de-identification, and weighting. It contains the cleaned survey responses, standardized variable names and labels, survey-round identifiers, respondent survey identifiers, and the appropriate weight variables for analysis.

References

Cochran, W. G. (1977). *Sampling techniques* (3rd ed.). John Wiley & Sons.

International Telecommunication Union. (2023). *Digital development dashboard: Lebanon*.
https://www.itu.int/en/ITU-D/Statistics/Documents/DDD/ddd_LBN.pdf

Pew Research Center. (2019). *Use of smartphones and social media is common across most emerging economies*. <https://www.pewresearch.org/internet/2019/03/07/use-of-smartphones-and-social-media-is-common-across-most-emerging-economies/>

United Nations Development Programme (UNDP). (2020). *Nothing and everything to lose: Results from a qualitative WhatsApp survey of Palestinian camps and gatherings in Lebanon*.

United Nations Development Programme (UNDP). (2024). *What is up on WhatsApp? A study on the impact of WhatsApp on tensions in Lebanon*.
<https://www.undp.org/lebanon/publications/what-whatsapp-study-impact-whatsapp-tensions-lebanon/>

World Bank. (2024). *Lebanon Economic Monitor 2024*.
<https://www.worldbank.org/en/country/lebanon/publication/lebanon-economic-monitor>

Appendices

Appendix A: Sampling and Weighting Framework

A.1 Population Benchmarks

In the absence of a national population census, all sampling and weighting operations for the LPS are anchored to official population estimates published in the annual Statistical Bulletin of the Lebanese Ministry of Public Health (MoPH) and estimates produced by the Central Administration of Statistics (CAS). These sources provide the most current institutionally endorsed estimates of the Lebanese resident population and define the calibration benchmarks applied during weight construction.

Table A.1: Estimated Resident Population by Age Group, Nationality and Sex, 2024

Age group	Lebanese			Non-Lebanese			Total		
	Male	Female	Total	Male	Female	Total	Male	Female	Total
0–4	132,304	121,278	253,582	58,056	53,218	111,274	190,360	174,496	364,856
5–9	151,167	144,054	295,221	66,333	63,212	129,546	217,500	207,266	424,767
10–14	158,446	145,512	303,959	69,528	63,852	133,380	227,974	209,364	437,339
15–19	179,197	154,870	334,068	78,634	67,959	146,592	257,831	222,829	480,660
20–24	175,341	161,589	336,930	76,942	70,907	147,848	252,283	232,495	484,778
25–29	134,638	144,994	279,632	59,080	63,625	122,705	193,718	208,619	402,337
30–34	114,848	123,932	238,781	50,396	54,383	104,779	165,244	178,315	343,560
35–39	113,619	128,677	242,296	49,857	56,465	106,322	163,476	185,142	348,618
40–44	103,388	132,549	235,938	45,368	58,164	103,532	148,756	190,712	339,470
45–49	114,448	138,875	253,322	50,221	60,940	111,160	164,668	199,815	364,482
50–54	130,437	150,324	280,761	57,237	65,964	123,201	187,674	216,287	403,962
55–59	116,802	129,780	246,582	51,254	56,949	108,203	168,056	186,728	354,785
60–64	97,061	114,190	211,251	42,591	50,108	92,699	139,653	164,298	303,951
65–69	82,072	82,072	164,143	36,014	36,014	72,028	118,085	118,085	236,171
70–74	68,408	68,408	136,817	30,018	30,018	60,037	98,426	98,426	196,854
75–79	52,599	47,339	99,939	23,081	20,773	43,854	75,680	68,112	143,793
80–84	38,292	38,292	76,585	16,803	16,803	33,606	55,096	55,096	110,191
85+	32,926	21,950	54,876	14,448	9,632	24,080	47,374	31,582	78,956
Total	1,998,073	2,046,609	4,044,682	876,774	898,072	1,774,846	2,874,847	2,944,681	5,819,528

Note: Table reproduced from the Lebanese Ministry of Public Health 2024 Statistical Bulletin.

Source: Statistics Department, Ministry of Public Health.

Table A.1 presents the estimated resident population by age group, nationality (Lebanese and non-Lebanese), and sex at the national level, as reported in the MoPH 2024 Statistical

Bulletin. When available, disaggregated district (Mouhafaza or Cada) level population counts are used for weight calculation.⁵

A.2 Weighting Methodology

Survey weights are constructed to ensure that estimates derived from the LPS are representative of the target population and to adjust for unequal probabilities of selection, differential response rates, and sampling variability introduced through the RDD design.

Base Weights

Given the RDD approach within valid Lebanese mobile number ranges, all sampled mobile numbers are assumed to have equal probability of selection within the constructed sampling frame. The base weight is therefore defined as the inverse probability of selection:

$$w_i^{(0)} = \frac{N_F}{n_F}$$

where N_F is the total number of eligible numbers in the generated mobile-number frame and n_F is the number of sampled numbers.

Non-Response Adjustment

To correct for differential response patterns across demographic and geographic groups, base weights are adjusted within calibration cells defined by combinations of district, gender, nationality, and age group. Each calibration cell is indexed by:

$$c = (d, g, z, a)$$

where d denotes geographic district, g gender, z nationality, and a age group.

The cell-specific response rate is computed as:

$$RR_c = \frac{\sum_{i \in S_c} w_i^{(0)} R_i}{\sum_{i \in S_c} w_i^{(0)}}$$

where $R_i = 1$ if the interview is completed and 0 otherwise, and S_c denotes respondents in cell c . The non-response-adjusted weight is then:

⁵ Up to 2019, MoPH statistical bulletins published population counts at the Cada level. Since then, only national-level estimates have been made publicly available. Despite their coarser geographic resolution, the more recent estimates are preferred because they better reflect the substantial demographic shifts that followed the post-2019 crisis.

$$w_i^{(1)} = \frac{w_i^{(0)}}{RR_c}$$

This step ensures that underrepresented response groups receive proportionally higher weights to correct for differential participation.

Post-Stratification and Calibration

Final weights are calibrated to external population benchmarks obtained from the MoPH. Calibration is performed using the same set of demographic and geographic margins defined in c , ensuring alignment between weighted survey distributions and known population totals.

Let N_c denote the estimated population total in calibration cell c . The calibration factor is defined as:

$$G_c = \frac{N_c}{\sum_{i \in S_c} w_i^{(1)}}$$

The final survey weight is then defined as:

$$w_i = w_i^{(1)} \times G_c$$

This ensures that the weighted sample reproduces known population margins across key demographic and geographic dimensions.

Appendix B: Call Outcome Classification System

Data collectors record a standardized outcome code for every telephone number dialled. These codes are used to monitor fieldwork progress, calculate response rates, and construct non-response adjustment weights.

Table B.1: Call Outcome Classification

Code	Label	Definition
CI	Completed Interview	Respondent was reached, met all eligibility criteria, consented, and completed the interview.
PI	Partial Interview	Respondent consented and began the interview but did not reach the end. Partial interviews are reviewed to determine whether they meet the minimum completeness threshold for inclusion in the analytical dataset.
NC	Non-contact	The number was dialled but not answered, busy, or could not be connected. Numbers are called up to three times at different times of day before being classified as non-response.
RF	Refusal	The respondent answered but declined to participate or consented but terminated the interview early. Refusals are not re-contacted.
IN	Ineligible	The respondent was reached but did not meet the eligibility criteria (age under 21, non-resident, or otherwise excluded per Section 1).
NR	Non-response	The number was not successfully contacted after three call attempts and is removed from the active calling frame.

The survey response rate is calculated and reported using the AAPOR Response Rate 3 (RR3) standard definition for each survey round.