

# Future of Business Survey Methodology Note

JW Schneider

November 2, 2020

## **Abstract**

This document provides detail on the methods used to develop, field and weight the Future of Business Survey, run in collaboration with the Organization for Economic Co-operation and Development (OECD) and the World Bank. The document introduces the effort, then proceeds to descriptions of sampling, weighting and design of the instrument itself. The document includes a discussion of future work to be done to better understand the relationship between the sampling frame and wider global business population.

# 1 Introduction

## 1.1 Project Background

The Future of Business Survey (FOBS) is a collaboration between Facebook, the OECD, and the World Bank to survey businesses on Facebook on a recurring schedule and assess their challenges, opportunities and needs around the world.

This initiative began in 2016 when leaders at Facebook, the OECD, and the World Bank were approached by the Global Partnership for Sustainable Development Data about the possibility of working together to leverage the Facebook platform to survey small and medium-sized businesses (SMBs) around the world in a timely and cost-effective manner.

The goal of the Future of Business Survey is meant to complement traditional business survey data with near real-time information on the perspectives of online SMBs in more than 100 countries.

Many businesses are difficult to capture in traditional household surveys, such as micro and unincorporated businesses. These businesses are somewhat easier to survey through a broad social media platform such as Facebook.

In March of 2020, at the request of the OECD and the World Bank, Facebook adapted its standard bi-annual approach to run six monthly waves of the Future of Business Survey in order to provide regular information on the impact of COVID-19 on small businesses and their adaptation to the pandemic.

In order to achieve a sufficient monthly sample of the online SMB population in each country, these COVID-19 monthly waves sample 50 countries around the world where the Facebook SMB penetration is sufficient to achieve a meaningful sample.

It should be noted that the biannual surveys cover a larger set of countries due to being less constrained by sample, the details of which can be found in the appendix.

The first of these monthly surveys (Wave I) surveyed around 30,000 SMBs. Example findings include observing over 1 in 4 small businesses had closed during the pandemic. In addition, of those that were still operational, a third had reduced their workforces, and nearly two-thirds reported lower sales compared to the same period last year.

## 1.2 Independent research uses of the Future of Business Survey

The OECD, the World Bank, and other research institutions have used the Future of Business Survey for a number of independent research products, several of which we detail below.

The World Bank has developed a working paper on [“Tackling the Global Profitarchy,”](#) which discusses a hierarchy of earnings, with male-owned businesses in male-dominated sectors earning the most, women in male-dominated

sectors and men in female-concentrated sectors in the middle tier, and women in female-concentrated sectors at the bottom.

There is also a working paper on examining the “[COVID-19 Pandemic Through a Gender Lens](#)”, which examines differences in closure rates among female and male-owned SMBs using FOBS data

Similarly, the OECD has used FOBS data for a number of country-level factsheets (see example [here](#)) on differences between male- and female-owned entrepreneurs, as well in their annual [Entrepreneurship at a Glance](#) report and [SME and Entrepreneurship Outlook](#) report.

The country level aggregated dataset from the Future of Business Survey is provided for free and is available on the [World Bank Data Catalog as well as on Humanitarian Data Exchange](#).

This document outlines the processes involved in implementing this survey from the Facebook perspective.

## 2 Sampling and Fielding

### 2.1 Target Population

The target population of the Future of Business Survey are Facebook Page admins that are business owners and managers.

There are over [80 million businesses](#) estimated to be represented in this sampling frame.

A Facebook Page serves as a key interface for business activity on the platform and functions as the organizational equivalent of a profile or in some cases supplanting a business’s website.

We cannot infer from data available outside the survey itself whether or not respondents hold a leadership position in any enterprise or if they are associated with any enterprise. We do, however, know that the rate of identification as a business owner or manager is between 5 and 20 times higher among Page admins than among the general population at Facebook (ratio depends on country).

A Page is a near-prerequisite for a business to engage in advertising or generating content for Facebook audiences, and therefore the set of Page admins is highly likely to contain almost all businesses on the platform. Simultaneously, a page has very low barriers to adoption: having a page does not require a business to engage in either advertising or content generation at any point.

Though we include all respondents in our reported data, we advise restricting analysis to self-identified owners and managers of small and medium enterprises.

The probability of the same business being resampled in any given wave has been calculated to be less than 0.06% among this population, but including employees raises this probability slightly.

Business owners and managers are most plausibly in possession of relevant business knowledge, and this distinction is important for setting our survey

population. Unlike enterprise surveys conducted at the firm, we cannot in this survey count on the presence of nearby specialized roles to answer specific questions. The owner or manager of a smaller enterprise is more likely to have knowledge of business plans, finances, and history. This decision does, however, influence decisions made around specific elements of survey content.

## 2.2 Sampling

Given our defined target population (see above) Facebook implements a modified simple random sample for invitations within each country.

First, we define our sampling frame as all Facebook accounts that are admins of active Facebook Pages.

At Facebook, all accounts, including those of Page admins, are randomly assigned to general survey eligibility.

Surveys at Facebook are randomly sampled, though sampling frames may vary from study to study (for example, focusing on Page admins).

Once a respondent has been invited to a survey (not necessarily replied) they are ineligible for another survey for up to six months. This is to reduce response burden on our populations and maintain general population response rates and response quality. This is also intended to minimize potential population bias: ineligibility is based on selection for a survey, not response to a survey. This is also one of the reasons we field the survey biannually: to make sure the population from wave to wave is as consistent as possible.

Though Facebook can implement more advanced stratified sampling, we find this is unnecessary given our ability to sample broadly enough to ensure survey weights do not induce excessive design effects (see appendix). This sampling procedure does, however, call for the construction of proper survey weights, the process for which is outlined below. We weight back to the base population we aim to represent, much as is done with all online panels.

## 2.3 Fielding and quotas

Fielding proceeds until country-level start quotas are met or the pre-defined fielding window has passed (which is the end of the calendar month for the COVID-19 waves or two weeks after initial fielding for the bi-annual waves). Quotas are set to ensure about 400 business owners and managers replying to the survey in most of our countries. This number was selected to be consistent with the initial waves of the FOBS.

We set lower targets for countries where the population is not large enough to sustain 400 business owners and managers replying, but do not set any targets lower than 100 business owners and managers.

We cannot predict beforehand how many business owners and managers will reply given a certain number of responses, but we have estimates from prior surveys that dictate the initial sampling quotas. We do not sample for complete surveys, but rather started surveys (this discrepancy predictable and is accounted for when we set our country-level sampling quotas).

## 3 Weighting and representativeness

### 3.1 Survey non-response patterns

Because Facebook does not use a stratified sampling approach within a given country the sample is not guaranteed to be representative of the target population (i.e. small and medium-sized businesses with a Facebook Page) if a given characteristic is also correlated with survey non-response in the context of the sampling procedure noted above.

Facebook is, however, in a unique position of knowing a considerable amount of information about both respondents and non-respondents, and over time Facebook has learned which characteristics are most closely associated with non-response to Facebook surveys. These include, but are not limited to:

- App language settings
- Recent frequency of app use
- Time since joining Facebook
- Gender
- Age
- Approximate sub-national region
- Country
- Main mobile device operating system
- Main mobile device age

Indeed we do observe significant differences across these characteristics in terms of who is invited to take the survey and who ultimately responds. To facilitate analysis and ensure results are weighted to our population, we construct non-response weights for the entire sample.

Survey-level non-response weights are constructed using an inverse-probability weighting approach (IPW). Specifically, we use a logistic regression to predict the probability of response given eligibility for the survey (i.e. being in our target population). This is a one of several methods used for implementation of propensity weighting, and a key example from the literature can be found [here](#).

We favor this approach for a number of reasons:

- Interpretability: a parametrized model does lose some performance relative to non-parametric approaches, but is easily explained to users of these datasets and consistency can be easily maintained across sampling units (countries)

- Foreknowledge: we already have a very good sense of the mechanisms of survey non-response at Facebook, so we are able to specify models which are more likely to be close to the true underlying model of non-response behavior which helps to mitigate our performance tradeoff. Specifically, it [has been shown](#) that understanding of proper variables is possibly more critical than the specific method when generating weights for on-line survey samples
- Efficiency: using this approach weights can be constructed in a matter of hours, as opposed to the days it may take to tune more elaborate methods, and converges very reliably from wave to wave
- Consistency: this method leads to very stable weighting schemes from wave to wave, and does not require fine tuning of models between waves that might introduce researcher-associated variability in survey metrics over time
- Outcomes: this method tends to keep weights from being too inflated for any given respondent, which is a risk for some other methods

The model is fully interacted with country: each country gets its own model due to substantial heterogeneity in the correlation between variables and response rates across countries. The model is fixed to have the same number of parameters for each country, and to have the same number of parameters for each feature across countries (as opposed to using LASSO or elastic net approaches which may introduce variability in this value).

For example, in each country we code parameters for the top two languages, and allocate all other languages to an “other language” parameter. This is done for parsimony, cross-national model stability, and interpretability.

Initial weights are constructed using the inverse of the predicted probability of responding to the survey for every respondent. Final weights are then created by trimming the initial weights. Weights are trimmed at the 1st and 99th percentiles for each country, then rescaled to sum to the total number of respondents for each country. This approach prioritizes adjusting for non-response over fitting to representativeness strictly on any specific subset of characteristics. Still, we do generally observe much better fits to target population characteristics after weights are applied.

Though in particular countries we have a good sense of what the demographics of non-Facebook business leaders might be by industry, gender, or age (in the US, for example) we do not have this data in most countries and in countries where such data are present they are often not comparable across countries.

This presents specific challenges for establishing representativeness of this survey to a baseline business population.

Rather than address this directly we, in agreement with the OECD and World Bank, take the approach of stressing that the data are intended to be representative of the Facebook Page admin population, rather than any national business population.

We take great care to avoid making any statements not supported by the data, and avoid making generalizations to national business environments or characteristics as a result

The resultant weights have desirable properties. The range of weights is generally in a narrow, acceptable range both across and within countries (see appendix). [Design effects](#) from unequal weighting are generally about 2 or less (with the exception of Saudi Arabia), which places them within acceptable survey bounds (see appendix). The log of weights are well-balanced across the sample, suggesting an acceptable distribution and not indicative of severe estimation issues (see appendix).

### 3.2 Item non-response patterns

Within a survey we also observe non-response to specific items, including our observed non-response to the survey overall. These patterns are a function of the length of the survey, logical gating, and question format.

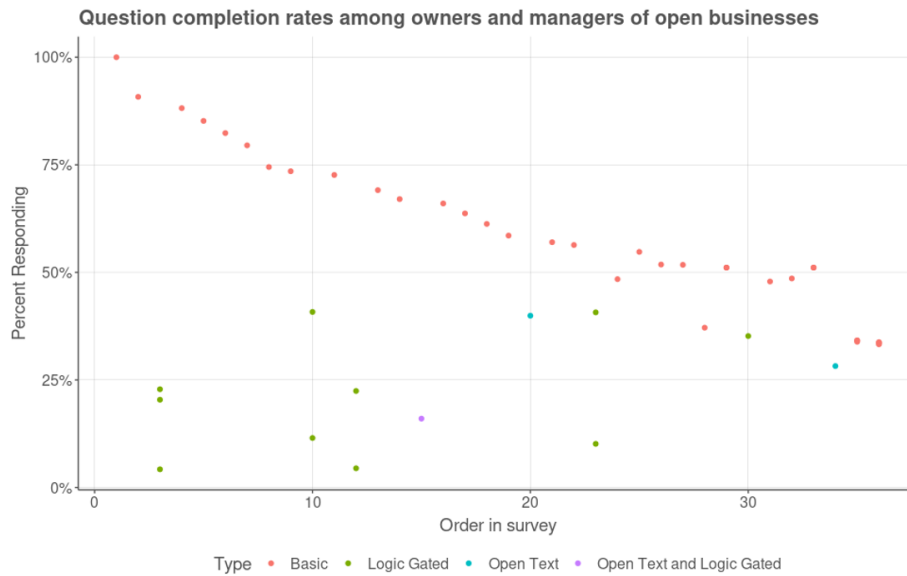


Figure 1: Percentage of survey responses with non-empty value by question order and type for the July wave of the FOBS COVID-19 survey

Furthermore, these non-responses are somewhat-correlated with characteristics, but this relationship is not strongly associated with survey length or question type after the first few items. Throughout the survey, non-respondents tend to be those who are also less likely to respond to the survey in general.

Below we demonstrate this phenomenon by comparing the sum of weights to the count of responses for each question. As the sum of weights is consistently below that of the count of respondents we can infer the average respon-

dent has a weight less than 1, implying they were more likely than average to respond to the survey generally. This gap is fairly consistent throughout the survey, but is slightly larger toward the beginning of the survey.

This tells us many of these respondents are breaking off from the survey fairly early and replying to very few questions. Those that stay with the survey after the first few pages appear to be skipping at similar rates as measurable on variables associated with survey non-response. We can therefore infer that item-level weights, which are already not common practice in survey analysis, would likely have a very small impact on results.

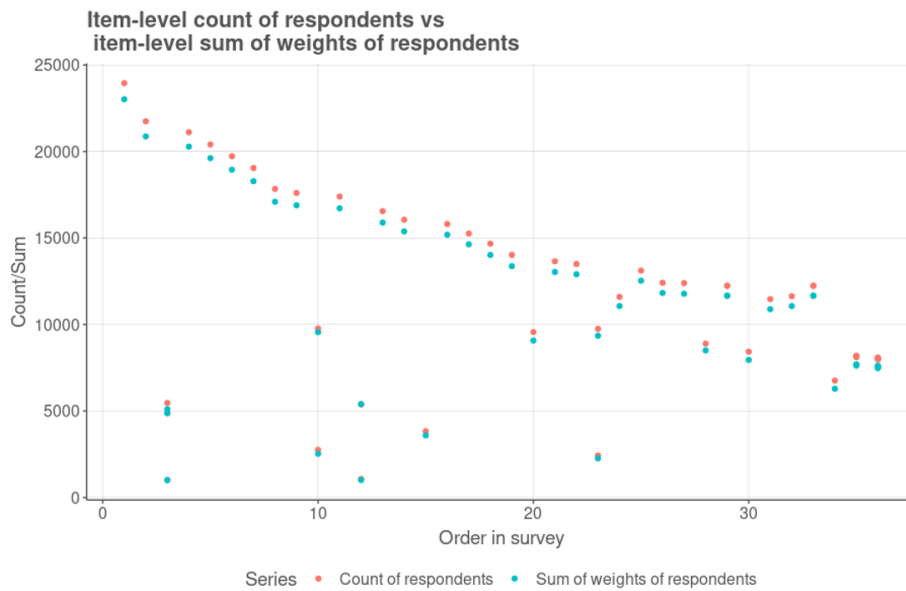


Figure 2: Sum of weights by question for respondents and non-respondents

### 3.3 Future representativeness work

Facebook may have future opportunities to better validate the potential representativeness of our survey population.

Facebook is performing internal comparisons of our survey data with nationally-representative survey data across multiple countries and contexts. This is occurring both totally internal to Facebook and also in partnership with outside organizations, and involves fielding surveys with the same or similar content, then comparing results both before and after the application of survey weights.

When possible we try to field simultaneously to generate fair comparisons. We hope to report on these efforts in the future following their completion. Further steps could be taken to validate representativeness, and possible future studies include but are not limited to:



- Comparison to or simultaneous fielding with existing planned household surveys run by partner organizations
- Exploration of differences in ability to represent populations in countries with lower internet penetration
- Understanding our ability to represent specific population such as business owners by comparing survey results with outside surveys of businesses

## 4 Instrument and Mode

### 4.1 Instrument

As is true with every Future of Business Survey instrument, all questions are drafted and prioritized by the three organizations. Akin to other FOBS questionnaires, the COVID-19 waves included questions on business sector, age of the business, gender of the business owner, and number of employees. Additionally, the latest waves included questions about closures, reduction in revenues, reduction of employees and challenges/needs of the business.

Facebook works with the OECD and World Bank to adapt the survey to the specific best practices Facebook has designed for surveys being conducted on its platform. Key principles revolve primarily around designing for a mobile audience, as most survey responses are collected on mobile devices.

Though no formal guide exists, [here](#) is an outline of many best practices at Facebook. It should be noted that, additionally, our own internal research strongly favors pagination over scrolling to increase completion rates.

Other core principles include:

- Simple questions that translate as neatly as possible to all languages the survey is fielded in
- Avoiding text entry due to higher cognitive burden and more challenging response-sharing problems associated with these question types
- Disallowing required questions and planning around question skipping behavior
- Avoiding complex question formats such as ranking and minimizing the use of matrix-type questions at the [recommendation](#) of the literature
- Avoiding agree/disagree scales due to [excessive cognitive burden](#) and other undesirable properties

We also make sure our survey is appropriate to our mode and business audience. Critically, we cannot count on businesses having all the same information handy when answering the survey that in-person interviews can count on. We cannot guarantee that our respondents will see the survey in a work

context, and have to adjust our expectations accordingly on both content and time to complete.

We also know very little about respondents, demographically, before fielding the survey. Therefore we must confirm many basic demographics in the survey instrument itself.

The instrument is translated into a set of languages appropriate to the countries being surveyed, restricted to those languages for which we have sufficient language support. We collect the list of most commonly used languages in-app for each country and strive to translate for at least 95% coverage of our target population. This is easily achieved in all target countries, though a small number of countries are not open to fielding due to our limited ability to support translations in certain languages (e.g. Armenia). Regardless of country we allow the survey to be taken in any translated language. Translations are completed by professional translators.

## 4.2 Survey Mode

Regardless of the wave, the Future of Business Survey is fielded on the Facebook platform directly. Potential respondents see an invitation at the top of their Facebook News Feed upon opening the app or website. They must click through the invitation to be taken to the introductory page of the survey itself.

We adhere to many practices recognized [in the literature](#) as effective for this mode. An example invitation is shown below.

Respondents may be invited to the survey every 3 days if they have not responded, until fielding closes. However, respondents can only respond to the survey once and the system will stop showing the invitation after a person starts the survey. All responses are optional and in addition to the survey invitation in the app the survey introductory page has a paragraph explaining to respondents the context of their participation.

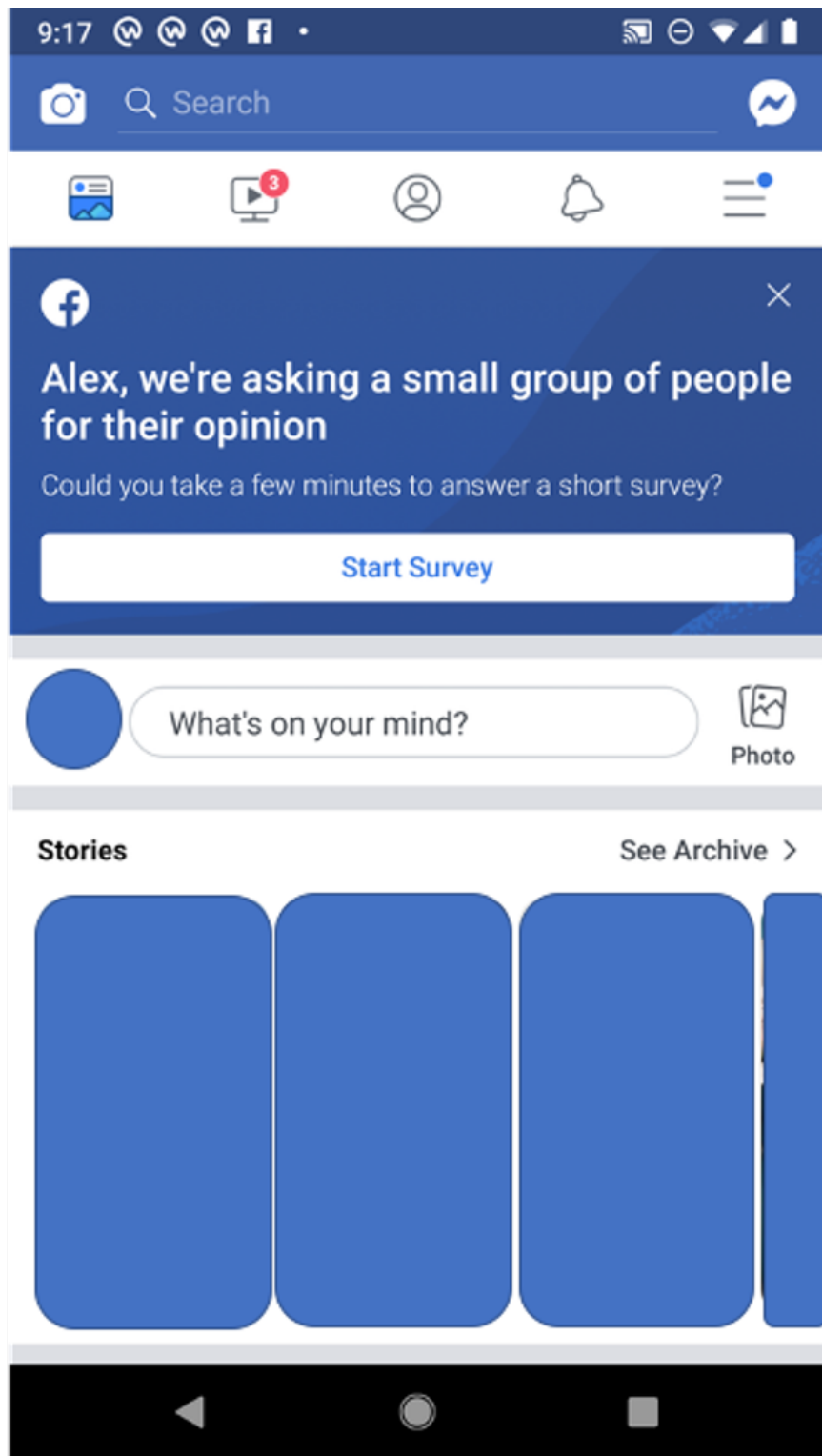


Figure 3: Example of invitation screen on a mobile device (Android operating system using the Facebook mobile app)

## **5 Conclusion**

This document provides an outline of the main methodological steps and innovations involved in conducting the FOBS. This information is provided for clarity and transparency in this project. Where appropriate we have referenced relevant literature and provided further quantitative analysis of survey outcomes.

We hope to build on elements of this document in the future, particularly regarding addressing how representative the Facebook survey population is of general business populations worldwide.

## 6 Appendix

### 6.1 Country Lists

Country	In COVID-19 waves	In bi-annual FOBS
UAE	Yes	Yes
Albania	No	Yes
Angola	No	Yes
Argentina	Yes	Yes
Austria	No	Yes
Australia	Yes	Yes
Azerbaijan	No	Yes
Bosnia	No	Yes
Bangladesh	Yes	Yes
Belgium	Yes	Yes
Burkina Faso	No	Yes
Bulgaria	No	Yes
Benin	No	Yes
Bolivia	No	Yes
Brazil	Yes	Yes
Botswana	No	Yes
Belarus	No	Yes
Canada	Yes	Yes
Switzerland	Yes	Yes
Côte d'Ivoire	No	Yes
Chile	No	Yes
Cameroon	No	Yes
Colombia	Yes	Yes
Croatia	No	Yes
Cyprus	No	Yes
Czech Republic	Yes	Yes
Germany	Yes	Yes
Denmark	Yes	Yes
Dominican Republic	No	Yes
Algeria	No	Yes
Ecuador	Yes	Yes
Estonia	No	Yes
Egypt	Yes	Yes
Spain	Yes	Yes
Ethiopia	No	Yes
Finland	No	Yes
France	Yes	Yes
UK	Yes	Yes
Ghana	Yes	Yes
Greece	Yes	Yes

Guatemala	No	Yes
Hong Kong	No	Yes
Honduras	No	Yes
Croatia	No	Yes
Haiti	No	Yes
Hungary	Yes	Yes
Indonesia	Yes	Yes
Ireland	No	Yes
Israel	Yes	Yes
India	Yes	Yes
Iraq	Yes	Yes
Italy	Yes	Yes
Jamaica	No	Yes
Jordan	No	Yes
Japan	Yes	Yes
Kenya	Yes	Yes
Cambodia	Yes	Yes
S. Korea	No	Yes
Kuwait	No	Yes
Laos	No	Yes
Lebanon	No	Yes
Lithuania	No	Yes
Libya	No	Yes
Morocco	No	Yes
Macedonia	No	Yes
Myanmar	Yes	Yes
Mexico	Yes	Yes
Malaysia	Yes	Yes
Mozambique	No	Yes
Nigeria	Yes	Yes
Nicaragua	No	Yes
Netherlands	Yes	Yes
Norway	Yes	Yes
Nepal	No	Yes
New Zealand	No	Yes
Oman	No	Yes
Panama	No	Yes
Peru	Yes	Yes
Phillipines	No	Yes
Pakistan	Yes	Yes
Poland	Yes	Yes
Portugal	Yes	Yes
Paraguay	No	Yes
Qatar	No	Yes
Romania	Yes	Yes
Serbia	No	Yes

Russia	Yes	Yes
Saudi Arabia	Yes	Yes
Sweden	Yes	Yes
Singapore	No	Yes
Slovenia	No	Yes
Slovakia	No	Yes
Senegal	No	Yes
El Salvador	No	Yes
Thailand	Yes	Yes
Tunisia	No	Yes
Turkey	Yes	Yes
Trinidad and Tobago	No	Yes
Taiwan	Yes	Yes
Tanzania	No	Yes
Uganda	No	Yes
USA	Yes	Yes
Uruguay	No	Yes
Vietnam	Yes	Yes
South Africa	Yes	Yes
Zambia	No	Yes

Table 1: Note this list does not include countries covered in the special "Rest of Africa" sampling unit which includes the rest of Sub-Saharan Africa but does not receive country-specific weights or quotas

## 6.2 Weight Characteristics

Country	Respondents	Max trimmed weight	Min trimmed weight	Design effect due to unequal weighting
Argentina	1429	8.5	0.3	2.1
Australia	1044	6.5	0.2	1.7
Bangladesh	1430	7.2	0.2	2.0
Belgium	1058	3.1	0.2	1.3
Brazil	1460	6.7	0.2	1.9
Cambodia	1445	5.3	0.2	1.7
Canada	1319	4.5	0.2	1.5
Colombia	1467	8.0	0.4	1.9
Czech Republic	1086	2.8	0.2	1.2
Denmark	632	3.1	0.2	1.4
Ecuador	1485	4.8	0.3	1.4
Egypt	2856	4.8	0.3	1.5
France	1645	4.0	0.2	1.4
Germany	1258	6.4	0.1	2.0

Ghana	1454	3.1	0.4	1.3
Greece	1415	4.3	0.1	1.7
Hungary	986	4.8	0.2	1.6
India	3671	6.2	0.2	1.7
Indonesia	1420	6.5	0.3	1.8
Iraq	1436	6.9	0.2	1.7
Israel	1092	5.3	0.1	1.9
Italy	2847	7.5	0.2	1.8
Japan	1102	5.8	0.1	1.9
Kenya	1425	4.1	0.2	1.5
Malaysia	1416	9.8	0.3	2.4
Mexico	2861	5.3	0.3	1.4
Myanmar	1455	5.9	0.3	1.6
Netherlands	658	4.5	0.1	1.8
Nigeria	2820	4.6	0.2	1.5
Norway	192	4.7	0.1	1.8
Pakistan	1455	5.0	0.2	1.6
Peru	1457	4.7	0.4	1.4
Philippines	1413	5.4	0.4	1.5
Poland	1359	4.8	0.2	1.5
Portugal	1418	5.3	0.3	1.5
Romania	1443	5.4	0.4	1.4
Russia	614	4.9	0.1	2.1
Saudi Arabia	1462	11.7	0.3	3.7
South Africa	1408	6.5	0.3	1.9
South Korea	159	5.6	0.1	2.1
Spain	1326	8.8	0.1	2.6
Sweden	524	5.0	0.1	1.9
Switzerland	315	4.2	0.2	1.5
Taiwan	1097	4.1	0.1	1.6
Thailand	2621	5.1	0.2	1.5
Turkey	1428	8.5	0.2	2.3
UK	2775	6.3	0.2	1.7
United Arab Emirates	1428	5.8	0.3	1.6
USA	2614	4.3	0.2	1.5
Vietnam	2846	8.0	0.3	1.9
Total (Excluding Rest of Africa Sampling Unit)	75026	11.7	0.1	1.7



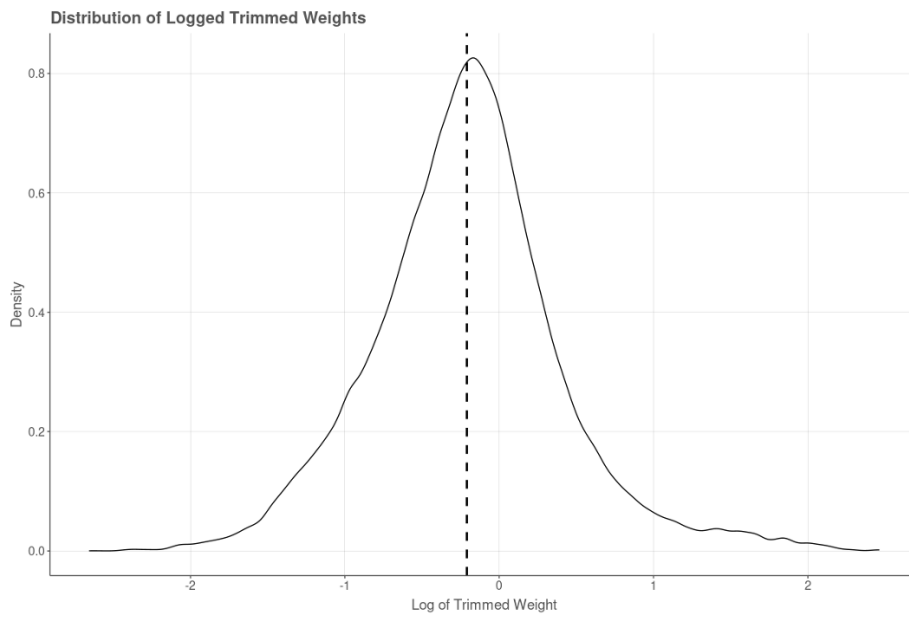


Figure 4: Distribution of weights