SAVING LIVES
CHANGING LIVES

Market Functionality Index

Sampling Guidance

WFP
World Food Programme

April 2020

MARKET FUNCTIONALITY INDEX

(*β-version*)

Sampling Guidance, April 2020

For additional information, please contact:

Research Assessment and Monitoring Division

wfp.economicanalysis@wfp.org

Oscar Maria Caccavale oscar.caccavale@wfp.org
Areum Han areum.han@wfp.org

Cover page photo credits: Oscar Maria Caccavale

World Food Programme

Via Cesare Giulio Viola, 68/70 - 00148 Rome - Italy

Contents

List of Figures

List of Tables

The **Market Functionality Index (MFI)** is a quantitative measure designed by WFP's Research, Assessment & Monitoring (RAM) and Supply Chain (SC) Divisions to benchmark market functionality along the following nine dimensions: 1) Assortment of essential goods, 2) Availability, 3) Price, 4) Resilience of supply chains, 5) Competition, 6) Infrastructure, 7) Services, 8) Food quality, and 9) Access and protection.

The MFI will be part of WFP's new Business Process Model for cash-based transfers. Specifically, it is expected to be used during the intervention design phase in the process 'Market assessment and risk identification', which is one of several feasibility assessments informing the choice of transfer modality. Additionally, it is supposed to be applied also during the delivery phase in the process of 'Market situation monitoring' to detect changes in market functionality over time.

The MFI (**α-*version***) was tested in Djibouti, Iraq, Mali, Zimbabwe, Haiti, and Malawi between August and October 2019.

The current version of the MFI (**β-*version***) is being tested in Bangladesh, Nigeria, Mauritania, Bolivia, Democratic Republic of Congo, Mozambique, Zambia, Iraq, and Syria between January and April 2020.

Following the COVID-19 crisis, a **reduced version** of the MFI administered with mobile calls has been developed and is being tested in Cameroon, Chad, Central African Republic, Iraq and Nigeria as of April 2020.

This document provides guidance on sampling for conducting a market assessment using the MFI tools. The first section introduces the concept of sampling. The second and third sections describe different sampling techniques (respectively probability and non-probability sampling) in the context of a market assessment. Section four provides practical suggestions for sampling focusing on the MFI, while section five describes all the resources available in the MFI package. Annex 1 displays the MFI questions, while Annex 2 concludes this guidance with the results from the Montecarlo simulation that was run to detect the ideal sample sizes in market assessments.

1. Introduction

WFP looks at markets at different levels and from various perspectives. All the activities that involve primary data collection and are conducted at the marketplace level are defined 'market assessment'. When conducting a market assessment, we call 'trader survey' the exercise of collecting primary data through an in-depth exercise that typically relies on enumerators (i.e. the persons employed to collect the data). These people are trained to carry out large data collection using a pre-defined questionnaire and following a survey plan. Alternatively, a rapid market assessment can be conducted through a checklist approach. In this case, an experienced market analyst gathers information from key informants. Combining both approaches is also common.

Before conducting a market assessment, it is important to distinguish between all the traders operating in the marketplace (i.e. the population in a market assessment) and the subset of these traders that are relevant for the research objective (i.e. the target population). While the two can coincide in relatively small markets, it is often the case that relying on a target population is a more practical way of tackling the assessment. There are several examples that we can highlight: if the market assessment is conducted to evaluate the capacity of the market to supply goods for more customers after the introduction of market-based programs, then the target population will be all the traders selling food that run medium to large scale businesses. Alternatively, if the objective is to investigate how WFP contracted shops perform, then these will become our target population. In case data collection is feasible only via mobile calls, then the target population will be those traders whose phone numbers are available. When the market assessment is conducted to investigate livelihood opportunities, then the target population will likely be petty traders. Alternative sampling frames[1] could also specify the eligible traders by predefining certain minimum qualifying criteria such as the amount of goods they normally sell, their shop size, or their willingness to provide reliable information. In the case of the Market Functionality Index (MFI), the target population will be traders selling cereals, other food, and non-food items.

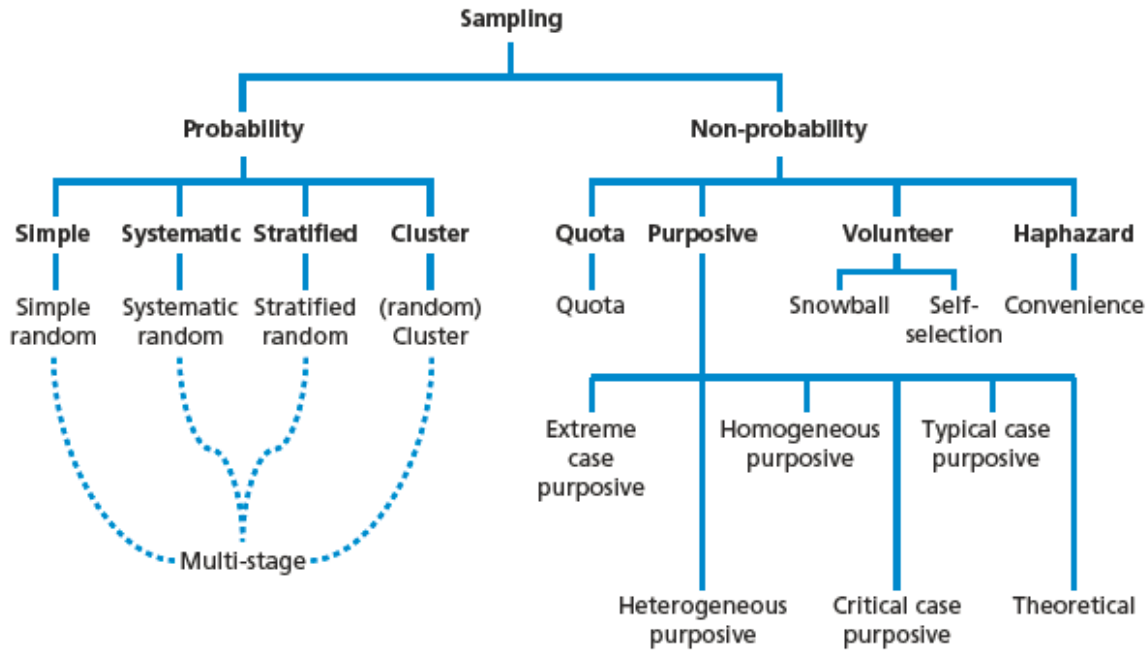
It is normally unpractical aiming to reach all members of a target population. Hence, unless we are assessing an extremely small marketplace, the analyst will have to survey only few members. The process of selecting a subset to estimate characteristics of the target population is called sampling.

The normative aspects of sampling fall into two main categories of survey approach: probability and non-probability methodologies. The difference between these depends on the procedures followed in the sample design and respondent selection.

[1] A sampling frame is a list of the items or people forming a population from which a sample is taken.

Figure 1 below provides an overview of these sampling techniques. We give a brief description of each methodology below and compare their respective advantages and disadvantages.[2]

Figure 1 - Sampling techniques



Source: Saunders, Lewies and Thornill (2016)

2. Probability sampling

Probability sampling methodology requires a clear understanding and definition of the target population, and a sampling frame like a census from which drawing the sample. Each case of the target population has equal and known chance of being selected. This makes the selected sample representative of its population, and therefore it is possible to make statistical inference[3] from the sample. There are several types of random sampling techniques, depending on how the population is structured. The most common ones are as follows:

- **Simple random sampling (SRS):** where each member of the population has an equal chance of being selected.

[2] A more comprehensive review is available in many books and is beyond the scope of this guidance. Among the others, see Fuller W.A., *Sampling Statistics*, John Wiley & Sons, Inc., New Jersey, 2009 and Saunders M., Lewis P. and Tornhill A., *Research Methods for Business Students*, Pearson Education Limited, Edinburgh. 2016, 7th edition, chapter 7.

[3] "The process of coming with conclusions about a population on the basis of data describing the sample is called statistics inference and allow you to calculate how probable it is that your result, given your sample size, could have been obtained by chance" (Saunders, Lewis and Thornhill, 2016).

- **Systematic random sampling:** this involves a random start and then selection of every *k*-th trader of the sample frame from then on, where k is equal to the population size divided by the sample size. In this case, if the target population is 100 and the sample size is 10, when the first individual selected is the 6th, then only the 16th, 26th,…, 96th respondents will be selected. This approach can be used in a relatively small marketplace.
- **Stratified random sampling:** when a heterogeneous population is split into fairly homogeneous groups/strata based on a factor that may influence the variable that is being measured; then SRS is applied in each stratum.
- **Cluster random sampling:** where the population is divided into similar groups/clusters. Each member of the population is assigned to just one cluster. A sample of clusters is chosen, using a probability method (often SRS). Only individuals within sampled clusters are surveyed. It may be more cost-effective to select respondents in clusters. Again, SRS or systematic random samplings are applied in each selected cluster.

These sampling procedures are used in specific circumstances depending on the homogeneity of the population and the inclusion of specific strata/groups/clusters with characteristics that are important to the researcher.

A truly random sample based on scientific principles produces unbiased and representative estimates of the population. Probability sampling is able to quantify the sampling error and thus puts the results in probability terms.

In an SRS, the desirable sampling size for a survey depends on the following parameters:

- Allowable *margin of error* (e), as no single sample can perfectly reproduce the entire population results, one has to allow a certain amount of error. Typically, the margin of error is between 1 and 10 percent.
- The *Z-score* corresponding to the desired confidence level. Most people use 1.64, 1.96 or 2.58 representing respectively the 90, 95, or 99 percent confidence level, taken from the standard normal distribution table.
- *Population size (N)*, e.g. the number of traders in the market (N), which determines a sample size that does not change substantially for populations larger than 10,000 individuals.
- The *response distribution* (p), which is conservatively set to 50 percent if no prior information exists about the population probability distribution.

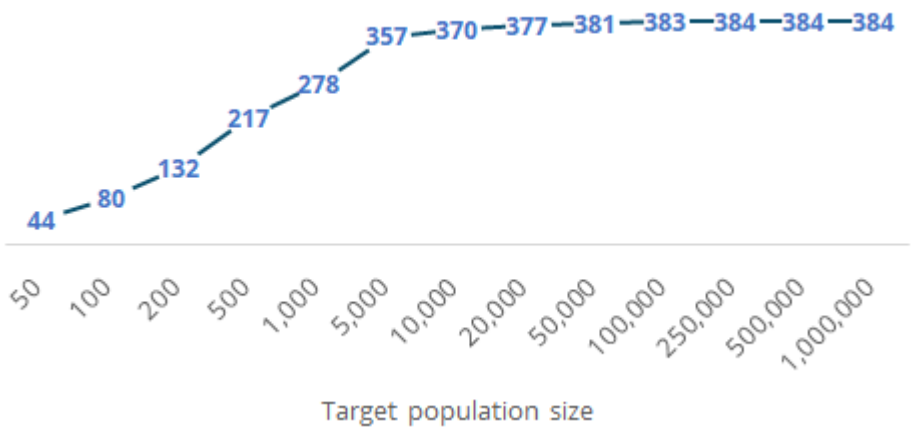
The formula to calculate the number of respondents is:

$$n = \frac{N}{1+\frac{(N-1)\cdot e^2}{p\cdot(1-p)\cdot Z^2}}$$

where n is the sample size.[4]

Figure 2 shows how the sample size changes as the target population grows.

Figure 2 - Simple random sampling with different target population size



Note: Confidence level 95%, Margin of error 5%, Response distribution 50%

Admittedly, for market assessments a sampling as such is not the ideal solution especially in relatively small marketplaces, where it makes little sense pursuing a probabilistic approach because the ultimate number of selected traders will be close to the entire target population. However, exceptions are possible as described in the Lebanon[5] and Palestine[6] market assessment reports.

3. Non-probability sampling

All the techniques described above have in common the feature that from a sampling frame it is possible to calculate the probability of each case to be sampled, and this probability is the same for each member of the population. Conversely, non-probability techniques are employed when either the probability of being selected is not known or when this is zero for some cases of the target population. They are normally used in market assessments for both practical and theoretical reasons. First, typically we have limited knowledge about the target population because a census of traders is not always available. Second, in market assessments we want to be sure that we assess relevant

[4] In the formula we use the finite population correction that is required to adjust the variance of the sampling distribution. This correction is not necessary when the population is extremely large compared to the sample size.

[5] The target population consisted in wholesalers selling food as per a 2010 census, and beverages and WFP contracted shops in Lebanon. Sampling was a one-stage stratified systematic random sample with respectively one and two strata (WFP, 2015 - page 20).

[6] The target population consisted in all retailers and wholesalers that sold food, beverages and household goods in the Gaza strip as per a 2012 census. Sampling was a one-stage stratified systematic random sample with three strata (WFP, 2017 - page 31).

marketplaces, and within those we aim to get a balanced view of the relevant actors in the business, without necessarily giving each trader a known and equal chance of being selected. Third, budget, time or security constraints may prevent large-scale data collection.

With non-probability techniques it is not possible to make statistical inference about the target population. However, if the goal is to get insights into market functionality and retrieve indicative information with no statistical grounds, then these qualitative methods of sample selection can be used, as they are more convenient and deliver critical information more quickly.

The non-probability methods are as follows:

Quota sampling: is used for large target populations and it assumes that the sample represents the target population. In fact, if we create groups that share common features relative to the proportions in which they occur in the target population, then variability of the different quota will be the same in the sample and in the target population. Calculations of quotas are based on relevant and available data, with the caveat that each group has sufficient respondents (at least 30) to enable meaningful statistical analysis (Saunders, Lewis and Thornhill, 2016). Quota sampling is similar to stratified sampling, except that respondents are not chosen randomly. In fact, enumerators select respondents based on who is easily accessible and apparently willing to answer to questions.

Purposive or judgment sampling: where observations are selected to fulfil a certain set of purposes defined by the analyst, such as the importance of the market (for market selection) in a given area, the size of traders' turnover and their reliability, trustworthiness and willingness to participate in the survey and provide information (for trader selection). Purposive sampling includes a judgment about what constitutes a suitable mix of people to participate in the sample. A fair representation of the population is made by deliberate decisions rather than a random process. Depending on the goal of the assessment, one can sample:

- **extreme cases**, based on the premise that some traders experiencing unusual or extreme outcomes may return significant insights into a specific phenomenon of interest for the market analysis;
- **heterogeneous cases**, if the goal is to capture the largest degree of variability from the sample, and hence grasping insights into market functioning from different angles and perspectives;
- **homogeneous cases**, which conversely shift the focus of the analysis on a specific subgroup of traders that share the same characteristics or play a unique role. For example, certain mid-size traders might be selected because they generally cater

to the poor, tribes or welfare recipients. There may be traders who are quite the opposite, dealing mostly with upscale clientele. Others could specialize in selling large volumes and thus offer the highest selling grade, quality or brand;

- **critical cases**, when selecting the most important traders or those that can share key information. This approach is pertinent when selecting 'big' traders based on the understanding that if they experience challenges (for example in the aftermath of a natural disaster or after the introduction of market-based interventions), then all the other smaller traders will face the same issues. For the same reason, as long as these traders have business connections outside of the marketplace, they will likely have better insights to offer to understand the 'big picture';
- **typical cases**, to provide an illustration of the 'normal' trader in a specific marketplace, using a list of criteria that define typical traders (to whom they sell, what they sell, what is their turnover, etc.). The idea is to cross-compare findings between samples, even though the typical case should not be misinterpreted as the average trader resulting from probability sampling;
- **theoretical sampling**: despite belonging to purposive sampling techniques, this is a rather different approach insofar as it involves collecting, coding and analyzing data in a simultaneous way to generate a theory. In a nutshell, the analyst will start collecting data, selecting respondents with a knowledge in the research area, then will analyze this information until further research ideas emerge, and then will go back to the field in search of further respondents that supposedly will bring new insights to the shaping research idea. The iterations of this approach can be multiple until the grounded theory becomes comprehensive. This was the approach used for developing the MFI survey through its α and β versions.

Volunteer sampling: these techniques involve volunteer participation of the respondents. The sample can be drawn using

- **snowball sampling**, when the enumerator contacts a trader, asking to further identify other cases, and so on. Typically, this approach can be used when the analyst wants to map backwards the supply chain from the retailer to different suppliers.
- **self-selection sampling**, when traders want to take part of a survey. This approach is mostly used during voucher-based interventions in the contracting phase, when traders answer positively to a call for expression of interest in doing business with WFP.

Haphazard or convenience sampling, when selection is based on certain convenience criteria, such as the ease of access to the marketplace or to the trader both geographically

and in terms of communication, or the availability of phone numbers in case of remote data collection.

Clearly, these methods suffer from selection bias and thus cannot be used if representativeness is important in the information obtained. Nonetheless, qualitative sampling methods are widely used as they typically rely on a much smaller sample size. They are cost efficient, faster to implement and can improve the consistency of estimates over time.

4. Sampling in Practice: Market Functionality Index

In practical terms, once the objective of the market assessment is defined, you might want to opt for a sampling approach and adjust it based on two key variables, namely (a) the number of markets to select and (b) the number of traders to interview per market. There is always a trade-off to take into consideration when either budget or time is a constraint: is it better selecting a higher number of markets with a lower number of traders per market aiming to have broader information coverage, or vice-versa, a lower number of markets and a higher number of traders aiming to have better data quality?

The rest of this section tries to iron out this conundrum, providing a practical guidance on the selection of markets and traders in the context of the MFI.

FIRST, SELECT THE MARKETS. For the purpose of this guidance, we define 'local markets' as the marketplaces that are directly used by a population of interest,[7] whereas regional/national 'market hubs' as those markets that play a regional or (trans) national role with an important wholesale function, hereby serving local markets within a trade corridor.

As for the local markets, if the population of interest resides in a number of geographical locations - assuming that each location has a different marketplace - then the assessment will be conducted in a certain percentage of those local markets. As a rule of thumb, we recommend a coverage ratio of between 25 and 50 percent (WFP, 2009). Therefore, if the assessment potentially covers a population residing in about 100 locations, then the number of selected marketplaces will span between 25 and 50.

When it comes to market hubs, it is recommended to include in the market assessment all key markets, including those in the capital and most relevant cities. The relevance of a market derives on overall volumes traded and their proximity to major ports or

[7] Note that the expression 'population of interest' is completely different from the expression 'target population' used before. In the context of a market assessment, the latter applies to traders, while the former applies to customers, more specifically to people vulnerable to food insecurity and hence potentially eligible to become WFP's beneficiaries.

neighboring countries. In case the number of relevant markets is too high, then it is a judgement call to reduce it according to the relevance within the trade corridor that serves the local markets.

For example, after the Bangladesh Cidr Cyclone, a trader survey was conducted sampling local markets in every other village (31 in total), as well as ten district level markets and four regional markets (WFP, 2007).

SECOND, SELECT THE TRADERS. The selection of traders strongly depends on the purpose of the market assessment.

If the goal of the market assessment is to score market functionality using the MFI, there are two cases to consider depending on whether the data is collected (i) through enumeration involving interviews with multiple traders in the market or (ii) from key informants. In the former case, we are in the context of a relatively in-depth trader survey[8] that can be administered either with face-to-face interviews (F2F MFI) or with remote calls (mobile MFI). In the latter case, we refer to a rapid market assessment where an experienced analyst is expected to fill out the MFI questionnaire after having discussed with key informants.

4.1. In-depth assessment using enumerators (F2F MFI):

IN THIS CASE THE ANALYST WANTS TO GET A BALANCED VIEW OF TRADERS, and they should opt for a purposive sampling where heterogeneous cases are selected based on the following criteria:

- Market chains: cereals, non-cereal food, and non-food items (NFI);
- Trader function: wholesale, retail, and collector;[9] and
- Trader volume size: small, medium and large.

How many traders do I need to select? If one relies on non-probability sampling, unfortunately, there is no magic number to suggest because guidelines "are virtually non-existent", even though the guiding principle should be the "concept of 'saturation', or the

[8] With in-depth trader survey we here refer to a data collection in a market assessment conducted using the MFI tool, irrespective of how the questions are asked (either with face-to-face interviews or with mobile calls). The MFI proposes a novel way of conducting trader surveys alternative to the existing WFP's approach described in the guidance 'How to conduct a trader survey' (WFP, 2009). In the humanitarian community there exist several alternative approaches to market assessment as described in this comparative table from the Cash Learning Partnership (CALP).

[9] Trader functions are defined as follows: "i) collector or assembler: person whose principal activity is to purchase from producers and sell to other traders; ii) retailer: person whose principal activity is to sell to final consumers; and iii) wholesaler: person whose principal activity is to purchase from traders, collectors/assemblers, and sell to traders" (WFP, 2009).

point at which no new information or themes are observed in the data" (Guest, Bunce and Johnson, 2006).

The typical approach used in market analysis suggests that traders should be chosen *per market* and *per product group*, including 2 retailers, 2 wholesalers and 2 collectors. As collectors often only exist for some products and during the harvest period only, they may not be available for an interview, in which case it is suggested to interview 2 wholesalers and 3 retailers at a bare minimum (WFP, 2009). Yet, this approach was developed when the focus of the market assessment was commodity-based, meaning that these numbers were to be multiplied by the number of relevant commodities.[10]

In the case of the MFI however, this rule of thumb seems extremely on the low side, because the focus of the analysis is not anymore on a specific commodity but rather on the overall supply of goods grouped in three market chain (cereals, non-cereal food and NFIs). In a more rigorous study that reviews the number of interviews used in quality assessments of different nature, Guest *et al.* (2006) suggest that 12 interviews are enough in case of homogeneous cases. However, if respondents are heterogeneous, the authors suggest to further increase the sample size to a number up to 30 (Guest, Bunce and Johnson, 2006).

We run a Montecarlo simulation to evaluate the effectiveness of different sample sizes and how they work in the context of a yes/no question[11], which is the case of the MFI. The Montecarlo simulation results confirm that some 30 traders would return a fairly good picture of the market. However, if one wants to compromise on the accuracy, then an acceptable window for the sample size remains between 7 and 12 retailers. We refer to the Annex for additional details on the Montecarlo simulation and the loss of accuracy when choosing different sample sizes.

Moving away from the Montecarlo simulation, these results can dramatically improve contingent on the interviewer's skills. In fact, the quality of the data is the "result of an interaction between the interviewer and the participant. There could be an argument, for example, which suggests that 10 interviews, conducted by an experienced interviewer will elicit richer data than 50 interviews by an inexperienced or novice interviewer" (Mason, 2010). As mentioned at the beginning of this section, the analyst needs to find a good compromise between sample size, the costs of the survey, and the skills of the interviewer(s).

What should be the right balance between different types of traders? There are three trader groups relevant for the MFI, namely cereals sellers, food sellers (beyond

[10] For example, with three relevant commodities (e.g. millet, sugar and vegetable oil) and assuming only retailers and wholesalers are available, then 9 retailers and 6 wholesalers per market should be interviewed.
[11] See Annex 1 for the full list of MFI questions.

cereals), and NFI sellers that should be equally represented in the sample. So, whatever number is decided for the sampling, you may think of replicating it for the three groups, if your budget allows that.

And how to deal with retailers and wholesalers (question X1 of the MFI)**?**

When it comes to **retailers**, these can then be chosen according to their business size. In the MFI we have some proxies for it, namely the approximate size of the shop (question X2), the number of counters (question X3), the number of employees (question X4), and potentially the number of stock-keeping units (SKUs) sold (question A2). For instance, in case you have opted for a sample size of 12, then you can consider dividing the sample in three retailer groups ('small', 'medium' and 'large' shops), and interviewing 4 per each typology.

For **wholesalers,** it makes sense to move away from a heterogeneous cases sampling and switch to critical cases instead. This typology of traders very likely will return a more homogeneous picture than retailers, thus their answers are expected to be more consistent.[12] Moreover, normally the number of wholesalers operating in many local (rural) markets is much below the 50 units of the Montecarlo simulation, hence allowing for a lower number of interviews.

In (rough) summary, traders to be interviewed per market should be:

Table 1 - Sample size summary

Reference	Retailers	Size			Wholesalers	Total per group	Total per market (total per group x3)
		Small	*Medium*	*Big*			
WFP, 2009	3	*1*	*1*	*1*	2	**5**	**15**
Montecarlo simulation *(lower case)*	7	*2*	*3*	*2*	3	**10**	**30**
Montecarlo simulation *(medium case)*	10	*3*	*4*	*3*	5	**15**	**45**
Montecarlo simulation *(upper case)*	12	*4*	*4*	*4*	7	**19**	**57**
Guest et al., 2006 *(homogeneous cases)*	12	*4*	*4*	*4*	12	**24**	**72**
Guest et al., 2006 *(heterogeneous cases)*	30	*10*	*10*	*10*	30	**60**	**180**

Note: The column 'Total per market (total per group x3)' refers to cereals, other foods and non-food items groups, and is the result of column 'Total per group' times three. The breakdown by size is indicative if one wants to have a balanced overview of traders in a marketplace

It is important to notice that we are talking about observations, meaning for example that if one trader sells both cereals and other foods, then it is possible to discount one from the suggested sample size. While potentially one can sample all traders that handle at least two of the three market chains relevant for the MFI to reduce the sample size, we would recommend to still consider a certain degree of specialization between traders in the sample.

[12] Recalling the Montecarlo simulation, dealing with wholesalers would probably imply having more homogeneous answers, hence allowing to reduce the sample size as per Table 2 to Table 4 in the Annex 2.

4.2. In-depth assessment using enumerators (Mobile MFI):

IN CASE THE MARKET ASSESSMENT IS ADMINISTERED WITH REMOTE CALLS instead of face-to-face interviews, a convenience sampling approach (perhaps combined with a snowball approach to expand the base of traders that can be reached over the phone) is the only viable solution. In this case, differentiating between types of traders, market chains, and retailer sizes might be complex, and the analyst should expect a much lower quality of data.

However, if the phone numbers available make a decent sampling frame, then it is possible to refer to Table 1.

4.3. Rapid assessment with no enumerators (Rapid MFI):

IN THE CASE OF RAPID OR LIGHTER ASSESSMENT, an expert analyst can interview a few traders in the most important trading hubs (e.g. two or three) and visit some local markets (e.g. five). In addition, it is recommended to interview one key informant in the market, which may be the Head of the market, a Government agent responsible for market monitoring, the President of the association of traders, or if all of these are unavailable, a large trader who is prepared to answer general questions on the market functionality. This way of conducting the market assessment can be combined with in-depth exercises described above (both F2F and mobile MFI), as it normally returns good background information that are critical for the analysis. The analyst may still use the MFI tool as a checklist, so that at the end of the rapid assessment it would be still possible to calculate the score for the marketplace.

5. MFI package

The MFI comes with a full package that includes:

- ✓ Technical Guidance
- ✓ Practical Guidance
- ✓ Sampling Guidance
- ✓ Questionnaires in XLS forms both for Face-to-face and Mobile data collection, *(available in English, Spanish, French, Arabic and Portuguese)*
- ✓ Video tutorial for enumerators for Mobile data collection (*available in English and French)*
- ✓ Feedback survey for enumerators and WFP staff in the field, to collate all the feedback of this beta version
- ✓ Data Library for storing raw data, script, processed data, report, meta-data, and terms-of-reference of the market assessment
- ✓ Data Bridge for data exchange *(forthcoming)*
- ✓ Market Assessment Database *(forthcoming)*
- ✓ Dashboard *(forthcoming)*

Reference

Guest, G., Bunce, A. & Johnson, L. 2006. How Many Interviews are Enough? An Experiment with Data Saturation and Variability. *Field Methods*, 18(1): 59–82. https://doi.org/10.1177/1525822X05279903

Mason, M. 2010. Sample size and saturation in PhD studies using qualitative interviews. *Forum Qualitative Sozialforschung*. https://doi.org/10.17169/fqs-11.3.1428

Saunders, M., Lewis, P. & Thornhill, A. 2016. *Research Methods for Business Students*. 7th edition. Pearson Education Limited. 272–315 pp. (also available at https://www.pearson.com/store/p/research-methods-for-business-students/P100001214517).

WFP. 2007. Food Security Assessment SIDR Cyclone – Bangladesh. (also available at https://fscluster.org/sites/default/files/documents/WFP_Cyclone_Sidr_food_security assessment_Mar08.pdf).

WFP. 2009. How to Conduct a Trader Survey? *Food Security Analysis*.

WFP. 2015. Exploring Food Assistance Programmes: Evidence for Lebanon. Rome20 pp. (also available at https://documents.wfp.org/stellent/groups/public/documents/ena/wfp278453.pdf?_ga=2.166772478.133963078.1586534362-2133877750.1542967394%0D).

WFP. 2017. Market Assessment in the Gaza Strip. Rome31 pp. (also available at https://documents.wfp.org/stellent/groups/public/documents/ena/wfp292313.pdf?_ga=2.155877115.133963078.1586534362-2133877750.1542967394).

Annex 1 - MFI questions

Note: Questions labelled with an *X.* in italic do not contribute to the calculation of the MFI.

X. Type of shop

X1. Which type(s) of customers does the shop serve?

X2. What is the approximate size of the shop?

X3. How many check-out points/ cashiers does the shop have (incl. any type of cash register, manual or automated)?

X4. How many employees are present in the shop?

A. Assortment of Essential Goods

A1. Which products are normally sold in your shop?

A2. What is the [maximum] number of distinct items on sale in this shop/any of the shops in this market?

B. Availability

B1. Are there products that are currently scarce in this market/shop?

B2. Are you/traders afraid of running out of stocks within one week from now?

C. Prices

C1. Are there products whose prices greatly increased in the last 1 month?

C2. If we ask you/the traders in this marketplace what the price will be in a week from now, would you/they get it right?

D. Resilience of Supply Chains

D1. Responsiveness of supply chains

D1.1 Considering your customers' current demand, would your current stocks last at least one week?

D1.2 If you place an order today, do you expect to receive your products within a week?

D2. Vulnerability to supply chain disruptions

D2.1 Are most of your suppliers geographically located in the same place?

D2.2 Do you currently have more than one supplier?

D2.3 Does your business mostly rely upon a single supplier?

E. Competition

E1. Are there less than five traders in the market (by product group)?

E2. Does one trader control the market (by product group)?

X5. How many traders operate in the market?

F. Infrastructure

F1. Which of the following best describes the majority of the shops in this market/this shop?

F2. Which of these features apply in/nearby this shop/market?

X6. Which are the structures that better/best describe this market/shop?

X7. Does the shop have a Business License/Legal registration?

G. Service

G1. Shopping: which of the following applies to this shop?

G2. Check-out: which of the following applies to this shop?

X8. Which forms of payment are accepted?

X9. Maturity: Which of the following applies to this shop?

H. Food Quality

H1. Is food in the shops protected from exposure to water, heat, direct sunlight, pests, chemicals, or other contaminants?

H2. Are fresh fruits and vegetables in the shops well-separated from raw meat, poultry, fish or seafood?

H3. Are raw meat, poultry, fish or seafood and dairy products in the shops stored and displayed in refrigerated units that are on and working?

H3.2 Is refrigeration in the shops always working?

H4. Are all goods in labelled containers not exceeding their “best-use-before”/ "use-by" date?

H5.1 Are processed pre-packaged foods in the shops intact and in properly labelled containers?

H5.2 Is food in the shops free of visible signs of spoilage and bad smells?

H5.3 Are food packages intact and free from signs of decay or damage?

X10. Does the shop undergo any health certification/ inspection on hygiene or HACCP annually?

I. Access & Protection

I1. How many of the following access issues are observed in this market?

X11. If there has been a natural disaster, which was it?

I2. How many of the following protection issues are observed in this market?

X12. If there are social barriers, what are they based on?

X13. If there are physical threats, who are under these threats?

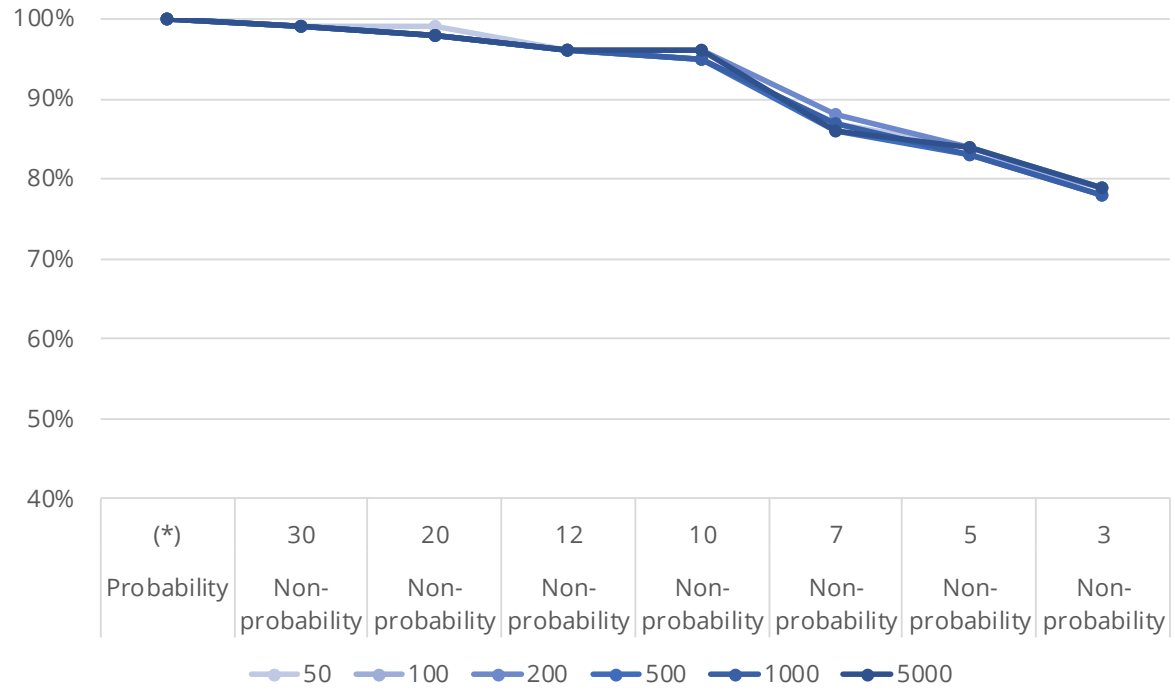
Annex 2 - Montecarlo simulation

We run a Montecarlo simulation to evaluate the effectiveness of different sample sizes and how they work in the context of a yes/no question, which is the case of the MFI. More specifically, we test how a sample of 3, 5, 7, 10, 12, 20, and 30 traders performs compared to a probability sample in markets of different sizes (i.e. with 50, 100, 200, 500, 1000 and 5000 traders). The Montecarlo simulation creates yes/no answers with a random Data Generation Process (DGP) that is iterated 10,000 times for each market size.

The prevalence of yes answers in the DGP can be 50, 60, 70, 80 and 90 percent. For example, if one takes question C1 of the MFI survey (*Are there products whose prices greatly increased in the last one month?),* then the case of 50 percent yes and 50 percent no answers implies that at the market level roughly half of the traders experience significant price increase and the other half reports the opposite. Admittedly, this is a rather hypothetical scenario that corresponds to the toss of a coin, because one would expect that prices increase for most of the traders at the same time. In the example above, we may better expect a 70:30 yes/no ratio if not more, depending on the severity of the price increase. All in all, based on the question and on the context (e.g. normal times, harvest time or lean season, sudden onset emergency, general inflation, depreciating currency, aftermath of a natural disaster, pandemic, etc.), we may expect that other DGPs are more pertinent than the 50:50 one.

With such a binomial variable, it is important that the sample correctly portrays the prevalence of yes/no of the actual data, because recalling question C1 of the MFI, we want to be able to say that prices have increased significantly or not at the marketplace level. This means that if the prevalence of yes is above (below) 50 percent, then the prevalence of yes in the sample needs to be above (below) 50 percent too. Figure 3 shows how many times out of the 10,000 iterations this is true. As expected, the accuracy drops the more we depart from a probability sampling, even though it remains above 90 percent until we reduce the sample size down to 7 respondents. These results are irrespective of the size of the market.

Figure 3 - Frequency out of 10,000 iterations when the prevalence of yes answers is either above or below 50 percent both in the sample and in the target population



Note: Sample size is shown in the horizontal axis, with (*) that refers to the probability sampling as per Figure 2. The lines in blue shades represent different market sizes as defined by the number of traders operating.

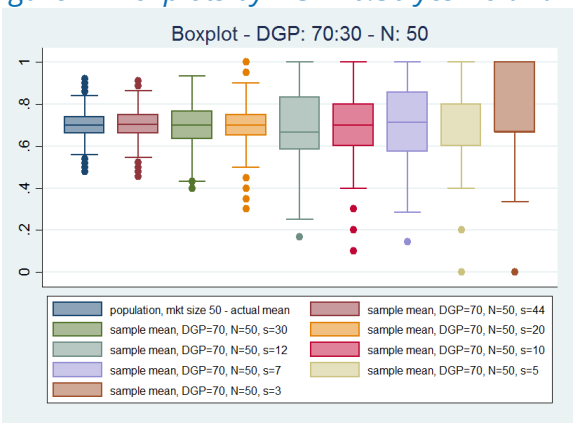
Note that the above results are smoothed down because we have considered the prevalence of yes/no answers and not if the sample mean is close enough to the population mean, i.e. the error of the sample. We show the error in percentage in Figure 5 to Figure 8,[13] noting that the more the population size grows, the more we would need to increase the sample size to reduce the errors. Two alternative measures of accuracy calculated from the error, i.e. the Root Mean Square Error (RMSE)[14] and the Mean Absolute Percentage Error (MAPE)[15] are shown in Table 3 and Table 4 respectively. The Montecarlo simulation results confirm that some 30 traders would return a fairly good picture of the market. However, if one wants to compromise on the accuracy, then an acceptable window for the sample size remains between 7 and 12 traders. Figure 4 shows how many times the average sample out of 10,000 iterations will be biased by outliers if the sample sizes is composed by 3 and 5 traders only, while it tends to remain consistent with all the remaining sample sizes, irrespective of the width of the box plots which of course become smaller as the sample size increases.

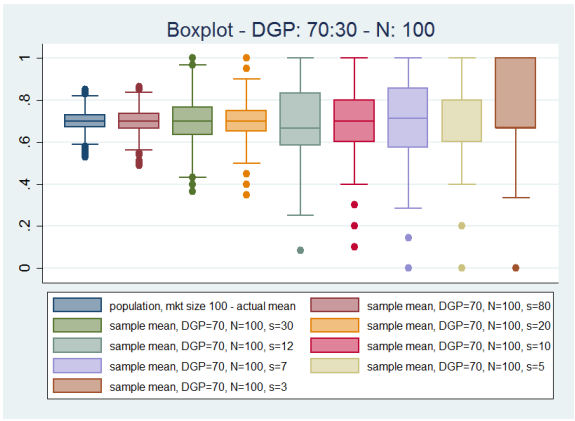
[13] For practical reasons, we show results only for DGP 70:30, and for population sizes of 50, 100, 200 and 500 traders. All the results of the Montecarlo simulation are available upon request at wfp.economicanalysis@wfp.org.

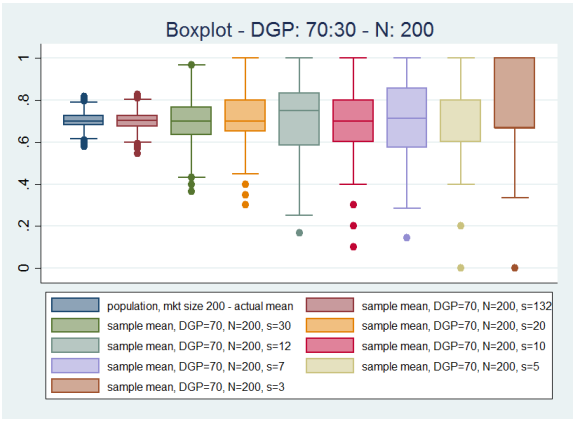
[14] RMSE represents the size of the "typical" error between the actual and sample averages. Considering the squaring process, it places a greater penalty on large errors. It is a statistic whose value determines the width of the confidence intervals for predictions.

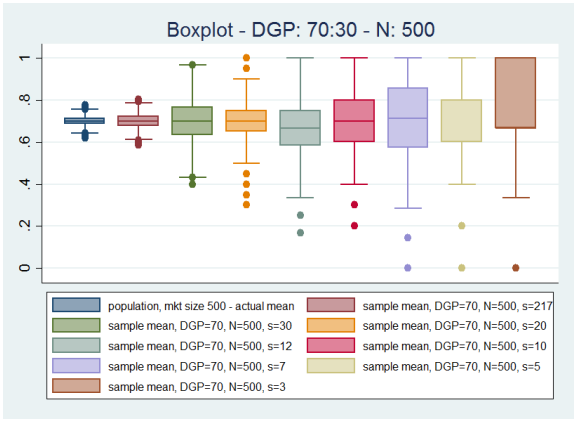
[15] MAPE is a statistic indicator in relative percentage. The closer MAPE approaches to zero, the better the fit of the sample. It can be useful for purposes of reporting, because it is expressed in generic percentage terms.

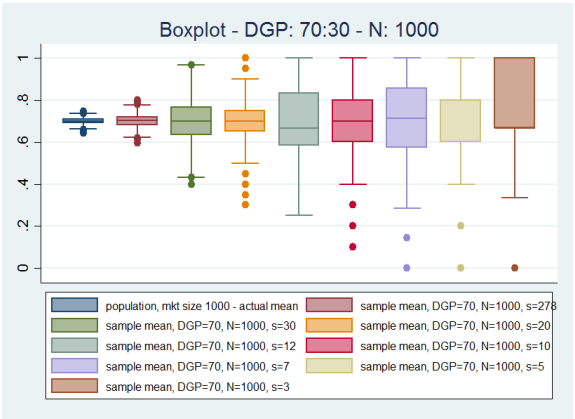
Figure 4 - Boxplots by DGP 70:30 yes/no and market size N











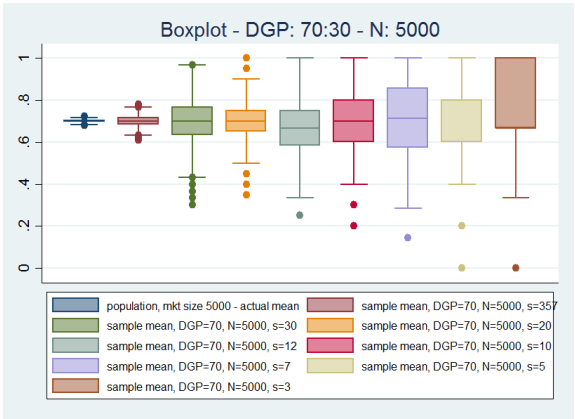


Table 2 - Number of correct hits by sample and market size

DGP - yes prevalence (%)	50	Number of traders in the market					
Sampling	Sample size	50	100	200	500	1000	5000
Probability	(*)	90%	86%	80%	73%	68%	59%
Non-probability	30	78%	68%	64%	58%	55%	53%
Non-probability	20	73%	65%	61%	57%	53%	53%
Non-probability	12	68%	62%	58%	56%	53%	53%
Non-probability	10	66%	61%	57%	54%	53%	52%
Non-probability	7	62%	59%	57%	53%	53%	52%
Non-probability	5	62%	57%	57%	54%	53%	51%
Non-probability	3	58%	56%	54%	52%	53%	51%

DGP - yes prevalence (%)	60	Number of traders in the market					
Sampling	Sample size	50	100	200	500	1000	5000
Probability	(*)	96%	98%	99%	100%	100%	100%
Non-probability	30	92%	90%	89%	91%	90%	91%
Non-probability	20	88%	87%	87%	87%	87%	87%
Non-probability	12	83%	83%	84%	84%	84%	85%
Non-probability	10	82%	82%	83%	84%	83%	85%
Non-probability	7	71%	69%	70%	71%	72%	73%
Non-probability	5	69%	67%	69%	68%	69%	69%
Non-probability	3	65%	63%	65%	64%	66%	66%

DGP - yes prevalence (%)	70	Number of traders in the market					
Sampling	Sample size	50	100	200	500	1000	5000
Probability	(*)	100%	100%	100%	100%	100%	100%
Non-probability	30	99%	99%	99%	99%	99%	99%
Non-probability	20	99%	98%	98%	98%	98%	98%
Non-probability	12	96%	96%	96%	96%	96%	96%
Non-probability	10	95%	95%	96%	95%	95%	96%
Non-probability	7	87%	87%	88%	86%	87%	86%
Non-probability	5	84%	83%	84%	83%	83%	84%
Non-probability	3	79%	79%	79%	78%	78%	79%

DGP - yes prevalence (%)	80	Number of traders in the market					
Sampling	Sample size	50	100	200	500	1000	5000
Probability	(*)	100%	100%	100%	100%	100%	100%
Non-probability	30	100%	100%	100%	100%	100%	100%
Non-probability	20	100%	100%	100%	100%	100%	100%
Non-probability	12	100%	100%	99%	100%	100%	100%
Non-probability	10	100%	99%	99%	99%	99%	99%
Non-probability	7	97%	96%	96%	97%	97%	96%
Non-probability	5	95%	94%	94%	95%	94%	94%
Non-probability	3	90%	90%	90%	90%	89%	89%

DGP - yes prevalence (%)	90	Number of traders in the market					
Sampling	Sample size	50	100	200	500	1000	5000
Probability	(*)	100%	100%	100%	100%	100%	100%
Non-probability	30	100%	100%	100%	100%	100%	100%
Non-probability	20	100%	100%	100%	100%	100%	100%
Non-probability	12	100%	100%	100%	100%	100%	100%
Non-probability	10	100%	100%	100%	100%	100%	100%
Non-probability	7	100%	100%	100%	100%	100%	100%
Non-probability	5	99%	99%	99%	99%	99%	99%
Non-probability	3	97%	98%	97%	98%	97%	97%

(*) Sample size in probability sample is 44, 80, 132, 217, 278, and 357, according to the population 50, 100, 200, 500, 1000, 5000. Confidence interval 95%, margin of error 5%.

Figure 5 - Errors in percentage, DGP 70:30 yes/no, Market size 50 traders

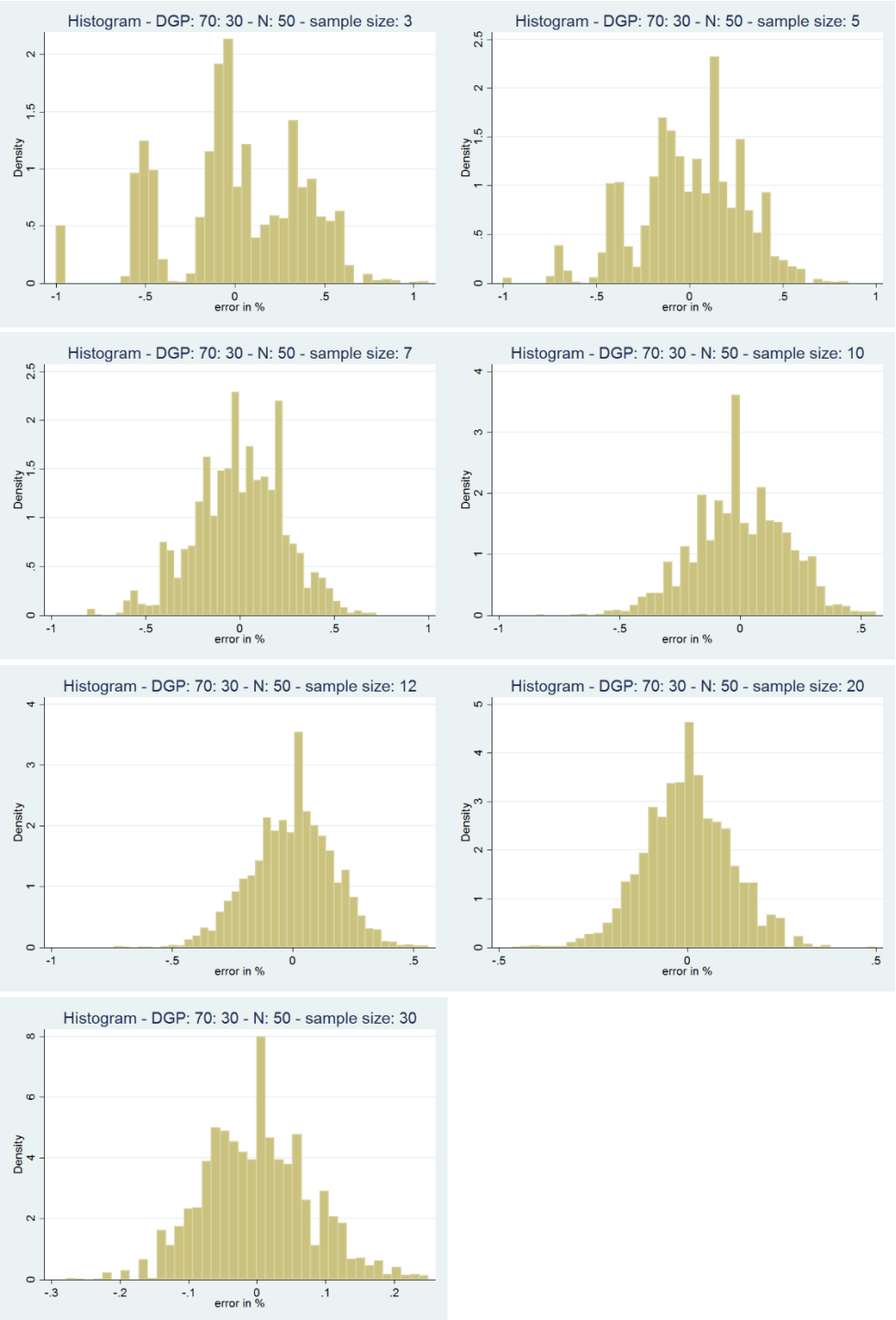


Figure 6 - Errors in percentage, DGP 70:30 yes/no, Market size 100 traders

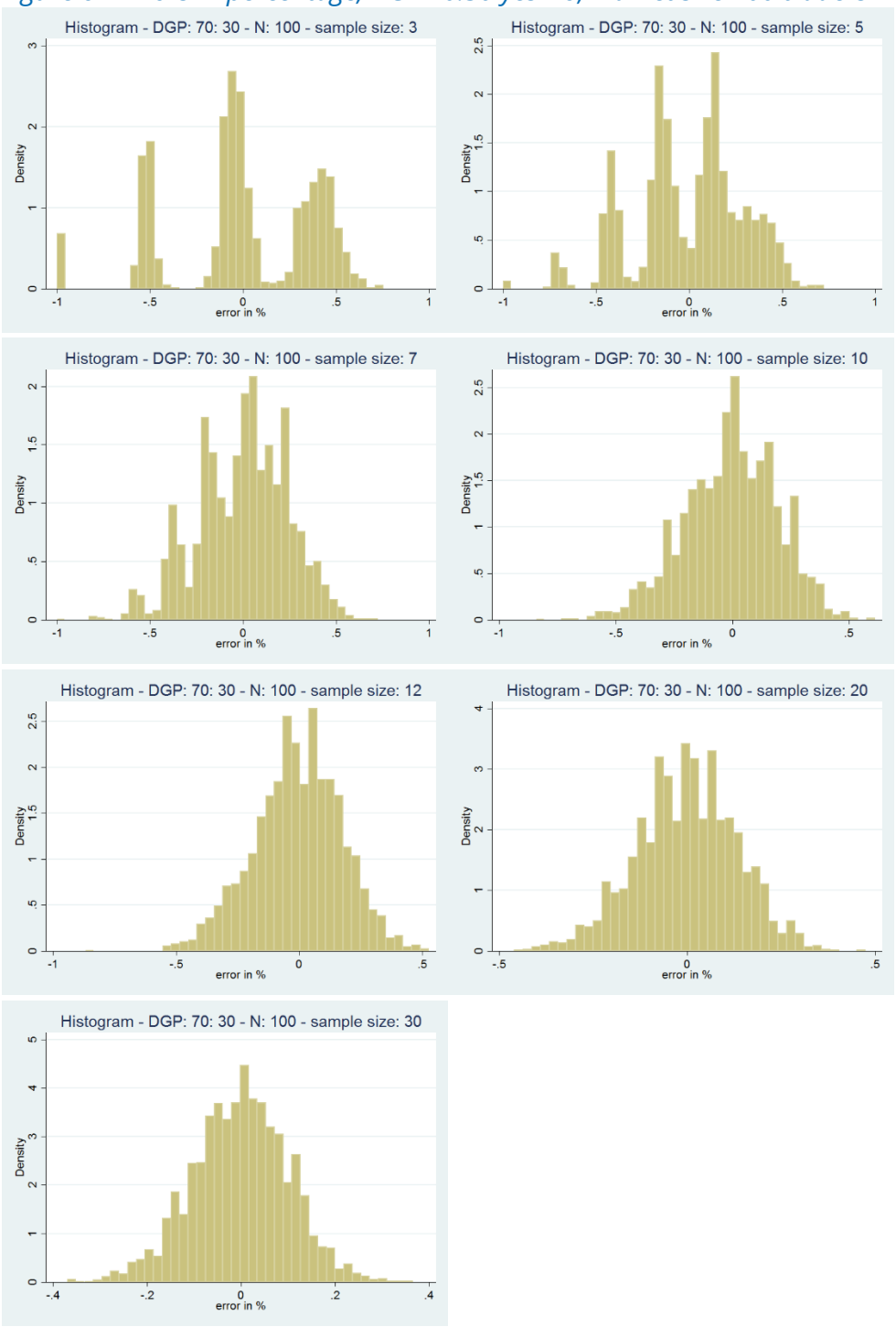


Figure 7 - Errors in percentage, DGP 70:30 yes/no, Market size 200 traders

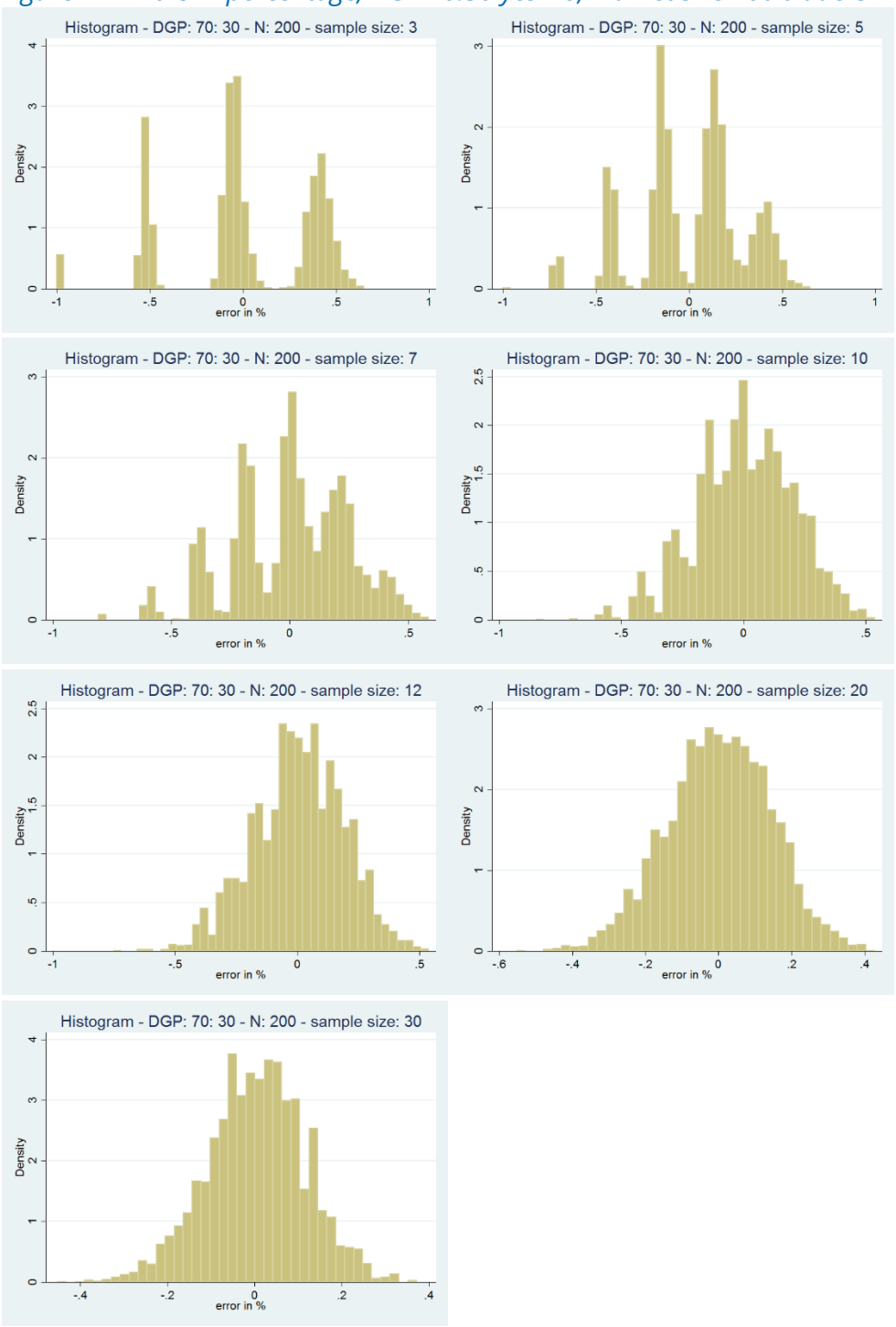


Figure 8 - Errors in percentage, DGP 70:30 yes/no, Market size 500 traders

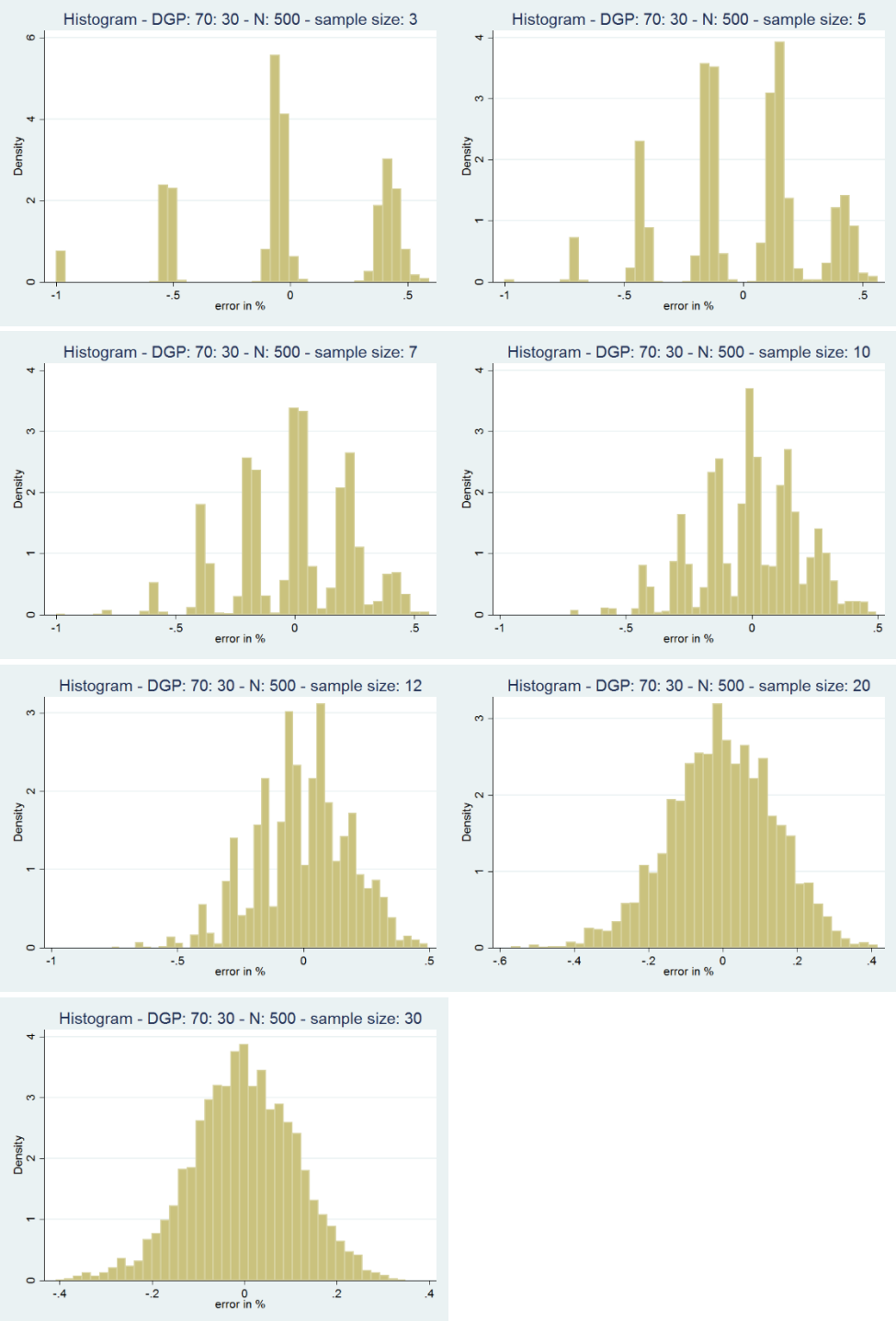


Table 3 - RMSE by sample and market size

DGP - yes prevalence (%)	50	Number of traders in the market					
Sampling	Sample size	50	100	200	500	1000	5000
Probability	(*)	2.10	1.95	2.01	2.03	2.00	2.02
Non-probability	30	4.68	6.04	6.55	6.98	7.22	7.24
Non-probability	20	6.97	7.93	8.34	8.63	8.90	8.74
Non-probability	12	10.11	10.80	11.16	11.33	11.22	11.30
Non-probability	10	11.35	11.94	12.08	12.47	12.15	12.48
Non-probability	7	14.05	14.72	14.69	15.24	15.25	15.82
Non-probability	5	17.27	17.84	17.77	18.43	18.62	18.69
Non-probability	3	23.26	24.26	24.35	24.95	25.34	24.96

DGP - yes prevalence (%)	60	Number of traders in the market					
Sampling	Sample size	50	100	200	500	1000	5000
Probability	(*)	2.08	1.95	1.98	2.02	2.01	2.03
Non-probability	30	4.53	5.98	6.68	6.73	7.10	7.06
Non-probability	20	6.80	7.91	8.23	8.41	8.73	8.68
Non-probability	12	9.88	10.80	10.94	10.96	11.51	10.88
Non-probability	10	11.20	12.00	12.12	11.92	12.51	11.52
Non-probability	7	13.92	14.76	14.89	14.45	15.16	14.11
Non-probability	5	16.62	17.54	17.28	16.87	17.22	16.12
Non-probability	3	22.41	22.73	23.18	22.96	23.47	22.50

DGP - yes prevalence (%)	70	Number of traders in the market					
Sampling	Sample size	50	100	200	500	1000	5000
Probability	(*)	1.89	1.82	1.86	1.85	1.86	1.91
Non-probability	30	4.18	5.60	6.20	6.23	6.65	6.63
Non-probability	20	6.19	7.39	7.91	7.89	8.07	7.89
Non-probability	12	9.18	9.77	10.27	10.43	10.64	10.52
Non-probability	10	10.34	10.96	11.17	11.24	11.31	10.97
Non-probability	7	12.83	13.42	13.46	13.61	13.54	13.39
Non-probability	5	15.68	16.73	16.67	16.97	17.46	17.04
Non-probability	3	20.25	20.70	20.35	20.50	20.46	20.03

DGP - yes prevalence (%)	80	Number of traders in the market					
Sampling	Sample size	50	100	200	500	1000	5000
Probability	(*)	1.69	1.60	1.60	1.65	1.60	1.65
Non-probability	30	3.62	5.00	5.43	5.64	5.80	5.70
Non-probability	20	5.40	6.54	6.84	6.88	7.21	6.91
Non-probability	12	7.87	8.80	9.14	9.15	9.40	9.25
Non-probability	10	8.95	9.81	9.96	9.80	9.89	9.53
Non-probability	7	11.19	11.85	12.21	12.29	12.61	12.64
Non-probability	5	13.53	13.66	13.71	13.41	13.57	13.31
Non-probability	3	19.28	19.86	20.05	20.20	20.56	20.41

DGP - yes prevalence (%)	90	Number of traders in the market					
Sampling	Sample size	50	100	200	500	1000	5000
Probability	(*)	1.30	1.20	1.19	1.23	1.23	1.21
Non-probability	30	2.76	3.69	3.89	4.27	4.20	4.21
Non-probability	20	4.08	4.83	5.09	5.26	5.20	5.12
Non-probability	12	5.96	6.47	6.69	6.71	6.85	6.85
Non-probability	10	6.58	7.16	7.34	7.10	7.21	7.02
Non-probability	7	8.48	8.94	9.23	9.30	9.61	9.59
Non-probability	5	10.68	11.04	11.49	11.62	11.83	11.80
Non-probability	3	13.70	14.02	14.16	14.30	14.48	14.52

(*) Sample size in probability sample is 44, 80, 132, 217, 278, and 357, according to the population 50, 100, 200, 500, 1000, 5000. Confidence interval 95%, margin of error 5%.

Table 4 - MAPE by sample and market size

DGP - yes prevalence (%)	50	Number of traders in the market					
Sampling	Sample size	50	100	200	500	1000	5000
Probability	(*)	4%	4%	4%	4%	4%	4%
Non-probability	30	10%	12%	13%	14%	14%	14%
Non-probability	20	14%	16%	17%	17%	18%	17%
Non-probability	12	21%	22%	22%	23%	22%	23%
Non-probability	10	23%	24%	24%	25%	24%	25%
Non-probability	7	29%	30%	29%	31%	31%	32%
Non-probability	5	35%	36%	36%	37%	37%	37%
Non-probability	3	47%	49%	49%	50%	51%	50%

DGP - yes prevalence (%)	60	Number of traders in the market					
Sampling	Sample size	50	100	200	500	1000	5000
Probability	(*)	4%	3%	3%	3%	3%	3%
Non-probability	30	8%	10%	11%	11%	12%	12%
Non-probability	20	12%	13%	14%	14%	15%	14%
Non-probability	12	17%	18%	18%	18%	19%	18%
Non-probability	10	19%	20%	20%	20%	21%	19%
Non-probability	7	24%	25%	25%	24%	25%	24%
Non-probability	5	28%	29%	29%	28%	29%	27%
Non-probability	3	38%	38%	39%	38%	39%	37%

DGP - yes prevalence (%)	70	Number of traders in the market					
Sampling	Sample size	50	100	200	500	1000	5000
Probability	(*)	3%	3%	3%	3%	3%	3%
Non-probability	30	6%	8%	9%	9%	10%	9%
Non-probability	20	9%	11%	11%	11%	12%	11%
Non-probability	12	13%	14%	15%	15%	15%	15%
Non-probability	10	15%	16%	16%	16%	16%	16%
Non-probability	7	19%	19%	19%	19%	19%	19%
Non-probability	5	23%	24%	24%	24%	25%	24%
Non-probability	3	29%	30%	29%	29%	29%	29%

DGP - yes prevalence (%)	80	Number of traders in the market					
Sampling	Sample size	50	100	200	500	1000	5000
Probability	(*)	2%	2%	2%	2%	2%	2%
Non-probability	30	5%	6%	7%	7%	7%	7%
Non-probability	20	7%	8%	9%	9%	9%	9%
Non-probability	12	10%	11%	11%	11%	12%	12%
Non-probability	10	11%	12%	12%	12%	12%	12%
Non-probability	7	14%	15%	15%	15%	16%	16%
Non-probability	5	17%	17%	17%	17%	17%	17%
Non-probability	3	24%	25%	25%	25%	26%	26%

DGP - yes prevalence (%)	90	Number of traders in the market					
Sampling	Sample size	50	100	200	500	1000	5000
Probability	(*)	1%	1%	1%	1%	1%	1%
Non-probability	30	3%	4%	4%	5%	5%	5%
Non-probability	20	5%	5%	6%	6%	6%	6%
Non-probability	12	7%	7%	7%	7%	8%	8%
Non-probability	10	7%	8%	8%	8%	8%	8%
Non-probability	7	10%	10%	10%	10%	11%	11%
Non-probability	5	12%	12%	13%	13%	13%	13%
Non-probability	3	15%	16%	16%	16%	16%	16%

(*) Sample size in probability sample is 44, 80, 132, 217, 278, and 357, according to the population 50, 100, 200, 500, 1000, 5000. Confidence interval 95%, margin of error 5%.

World Food Programme

Via Cesare Giulio Viola 68/70
00148 Rome, Italy
T +39 06 65131 wfp.org