

Reporting financial inclusion research results using **confidence intervals, margins of error and confidence levels**

October 2017

Established by



Authors

Grant Robertson and Bobby Berkowitz,
Data4FI, insight2impact

About insight2impact

Insight2impact | i2i is a resource centre that aims to catalyse the provision and use of data by private and public-sector actors to improve financial inclusion through evidence-based, data-driven policies and client-centric product design.

i2i is funded by the Bill & Melinda Gates Foundation in partnership with The MasterCard Foundation.

For more information:

Visit our website at www.i2ifacility.org.

Email Mari-Lise du Preez (i2i's Partnerships Manager) at mari-lise@i2ifacility.org.

Call us on **+27 21 913 9510**.



Contents

Introduction	2
Error reporting approach and tool	3
Competing estimates	3
Tracking change	3
Confidence intervals as the appropriate tool	3
Confidence intervals in practice	4
What to create confidence intervals for	4
Some important concepts	5
Caveats in interpreting the MoE	6
What affects the size of the MoE?	6
Calculating MoE on sub-groups	6
Practical application	7
Conclusion	10
Bibliography	11

Introduction

This note introduces a tool and approach to assist financial inclusion users in better understanding, interpreting and using their data. Adoption of this tool will enable the entire financial inclusion community to produce better research. This will increase credibility among users, as well as onlookers.

Financial inclusion users often have to choose between the reported figures of rival surveys. They also often want to understand whether and how the financial inclusion situation is changing. Users may take the figures at face value. Doing this may create confusion and lead to inappropriate interpretation and interventions.

“ Researchers need to ask whether the differing estimates from different surveys are really different and whether they are simply from an acceptable range of estimates. ”

A network diagram with various sized grey nodes connected by thin grey lines, forming a complex web-like structure that serves as a background for the top of the page.

Error reporting approach and tool

Competing estimates

Findex is a World Bank published survey of financial inclusion estimates in many countries around the world. There are other syndicated studies such as FII. Many countries also do their own bespoke demand-side surveys of financial inclusion.

Competing surveys lead to competing estimates of financial inclusion indicators. When these differ and users do not have the tools to compare the differences, users may lose trust. This undermines reporting on financial inclusion in the first place. The question researchers need to ask is whether the differing estimates from different surveys are really different and whether they are simply from an acceptable range of estimates.

Tracking change

Financial inclusion surveys are repeated to understand how key indicators move over time. This provides policymakers and other users with targets to set, as well as feedback on their initiatives. Users of this data look for trends in changes. These changes do not always move in the direction expected. It is important for users to understand that these unexpected changes may not be due to underlying phenomenon but are often due to the nature of survey data and implicit random error in survey estimates of underlying phenomena. The same principle applies here: researchers need to ask whether the differing estimates over time are really different and whether they are simply from an acceptable range of estimates. If the estimates are truly different, there will be no overlap in the confidence intervals.

Confidence intervals as the appropriate tool

There is a large debate in the scientific research community around the efficacy of significance testing and the over-reliance on the p-value to make claims about important scientific findings. There have been calls to scrap significance testing completely or to reduce the standard p-value from 0.05% to 0.01%. There have also been calls to consider alternative approaches such as reporting:

- Effect sizes
- Bayesian estimates
- Likelihoods
- Confidence intervals¹

We feel that the use of confidence intervals is most appropriate for financial inclusion surveys, as very little significance testing is conducted and results are usually presented as point estimates, which can be misleading. As such, this note will provide detail on this method.

When results of financial inclusion surveys are reported on, they often include statements like “55% of the population have a bank account”. An estimate from a survey like this is unlikely to exactly equal the true population quantity of interest due to sampling and non-sampling error. It is thus more accurate to think about an estimate as an interval as opposed to an exact point estimate.

1 <https://www.vox.com>



Confidence intervals in practice

In this note, we are interested in sampling error and how it can be quantified and used to report financial inclusion estimates more accurately. In a probability survey, every unit in the population has some known positive probability of being selected for the sample, and the probability of any particular sample being chosen can be calculated. It is this property that allows us to calculate the extent of sampling error present in our estimates. This estimate of the error is often called the Confidence Interval or Margin of Error (MoE). The Confidence Interval allows us to specify the intervals (or range) of values around an estimate with a certain level of confidence (usually 95%).

As an example, the result above should be reported on as 55% (+/-3%) of the population have a bank account. This simply means that, should we draw repeated samples from the same population, we are 95% certain that the true estimate of people with bank accounts will fall between 52% and 58%. An often overlooked fact is that there is still a 5% chance that the estimate will fall outside of this range, i.e. lower than 52% and higher than 58%.

What to create confidence intervals for

To keep things simple, only the most important indicators should be compared using confidence intervals. This will depend on the country but can be things like headline financial inclusion, as well as key interventions looked at. In addition, a clear explanation should be provided to each confidence interval done. This is to ensure that confidence intervals are not done as a simple box-ticking exercise but the purpose and meaning is clearly understood by a broad audience.

“ The Confidence Interval allows us to specify the intervals (or range) of values around an estimate with a certain level of confidence (usually 95%). ”



Some important concepts

Sampling error

Each time we draw a new sample from a population, the estimate of interest will differ from the true population value. This is because each new sample is different. The difference in estimates brought about by chance is called the sampling error.

Confidence Interval

Statisticians use a confidence interval to express the degree of uncertainty associated with a sample statistic. A confidence interval is an interval estimate combined with a probability statement (confidence level) (<http://stattrek.com/statistics/>). It basically is the range of possible estimates generated by an estimating process that would 95% of the time contain the true value of the parameter being estimated. The confidence interval is the range into which the true population parameter will fall, assuming a given confidence level (Malhotra & Birk, 2003). Confidence intervals are preferred to point estimates and to interval estimates, because only confidence intervals indicate (a) the precision of the estimate and (b) the uncertainty of the estimate.

Confidence Level

Some confidence intervals include the true population parameter, while others don't. In survey sampling, different samples can be randomly selected from the same population, and each sample can often produce a different confidence interval. Some of these samples will include the true population parameter and some will not (<http://stattrek.com/statistics/>). The confidence level is the probability that a confidence level will include the population parameter (Malhotra & Birks, 2003). Usually a 95% confidence interval is set, which means that 95% of the sampled confidence intervals will include the true value.

Margin of Error

The margin of error expresses the maximum expected difference between the true population parameter and a sample estimate of that parameter. To be meaningful, the margin of error should, like the confidence interval, be qualified by a confidence level. The margin of error is simply half the confidence interval. If the confidence interval is 10% with a confidence level of 95%, then the margin of error is +5% of the estimate and -5% of the estimate. For example, on a sample size of 400 with a confidence interval of 9.8% and a proportion of 50%, the MoE will be $\pm 4.9\%$. This means that the true value will fall between 45.1% and 54.9%, 95% of the time the population is sampled.



Caveats in interpreting the MoE

The importance of sample design

Demand-side surveys of financial inclusion typically involve idiosyncratic survey study designs. This happens whenever a survey uses existing population database references to “design” or create the survey’s structure, as well as “weight” or extrapolate the sample results to make broader population inferences. This matters because all decisions used to design a survey create their own kinds of error estimates. Any confidence interval needs to take the structure of error into account. Each survey has its own unique design and thus its own unique confidence interval calculation. Confidence intervals should therefore be individually calculated. There is simple computer software readily available for doing this, with the appropriate parameters to be identified and inputted.

It is important to note that the MoE does not take into account non-sampling related sources of error and non-response error and is thus by definition an underestimate of the true total error in a survey. All sources of potential error/bias should be considered when interpreting an estimate.

What affects the size of the MoE?

There are two key aspects that determine the margin of error:

1. The larger the sample the lower the MoE will be. It should be noted that an increase in sample size has diminishing returns for the MoE.
2. The more extreme the estimate the lower the MoE will be, e.g. a 10% estimate and 90% estimate have a lower MoE than a 50% estimate using the same sample size.

The absolute size of the population has very little to do with the MoE.

Calculating MoE on sub-groups

Margins of error typically are calculated and reported on for surveys overall but should also be calculated again when a subgroup of the sample is considered. Some surveys do not require every respondent to answer every question, and sometimes only certain demographic groups are analysed. In such cases the sample size and extremeness of the estimates will vary. The MoE will therefore also vary.

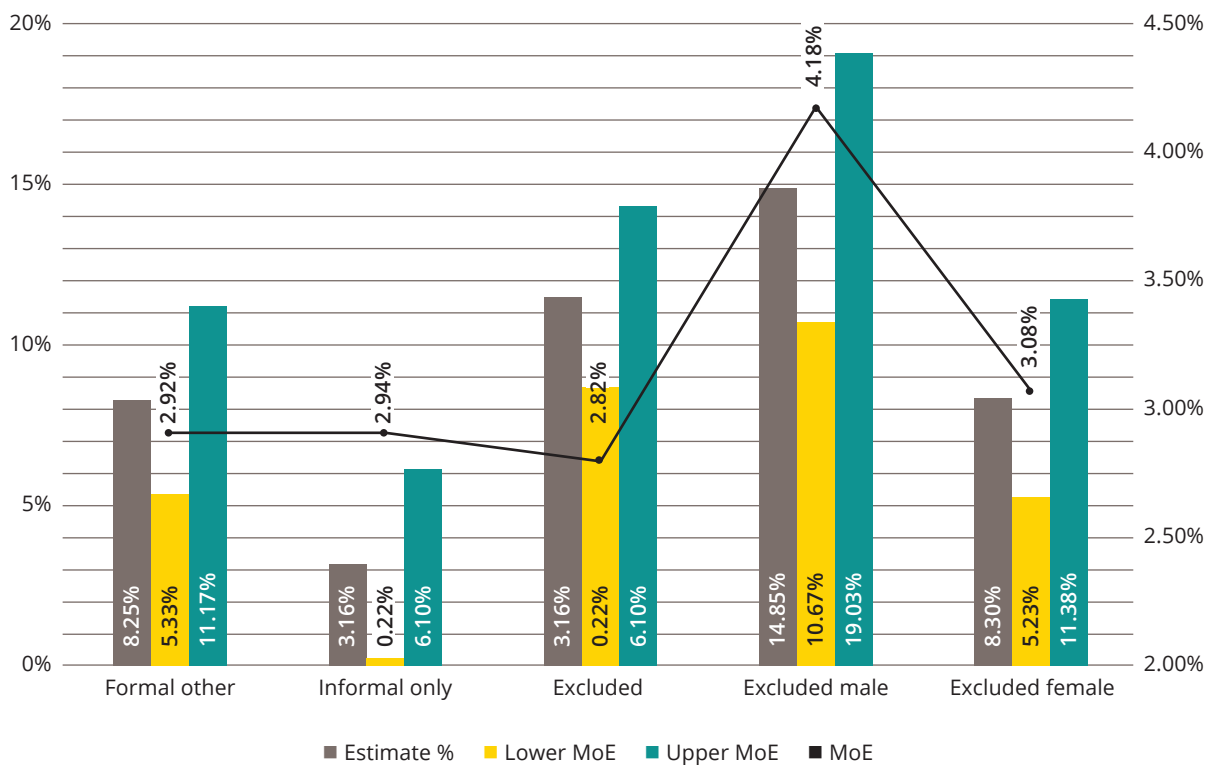
Practical application

In the charts below, the margin of error has been calculated individually for each FinScope Access Strand based on the realised sample size and actual proportion achieved. The blue bar is the actual estimate of the number of people in each strand, whereas the grey bar is this estimate plus the estimated margin of error and the orange block the estimate minus the MoE. The Formal strand has been excluded. If we look at the Formal Other strand, we can see that we are 95% confident that the true estimate falls between 5.33% and 11.17%.

If we take the excluded strands' results and break them down by gender, we can see that the MoE increases from 2.82% (both males and females included) to 4.18% (for males only) and the result for males being excluded thus ranges from as low as 10.67% to as high as 19.03%.

It is very important to communicate this band or interval continuity when citing estimates, especially when dealing with small sample sizes and proportions that centre around 50%, as that is where the MoE is the largest.

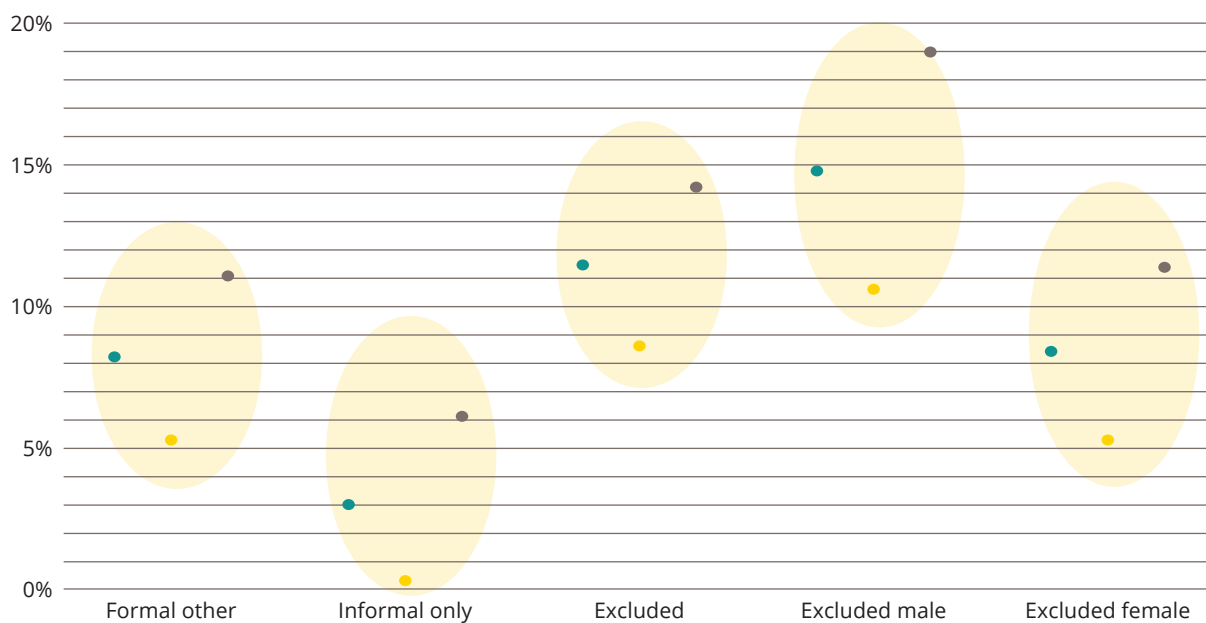
South Africa – Access Strand 2016



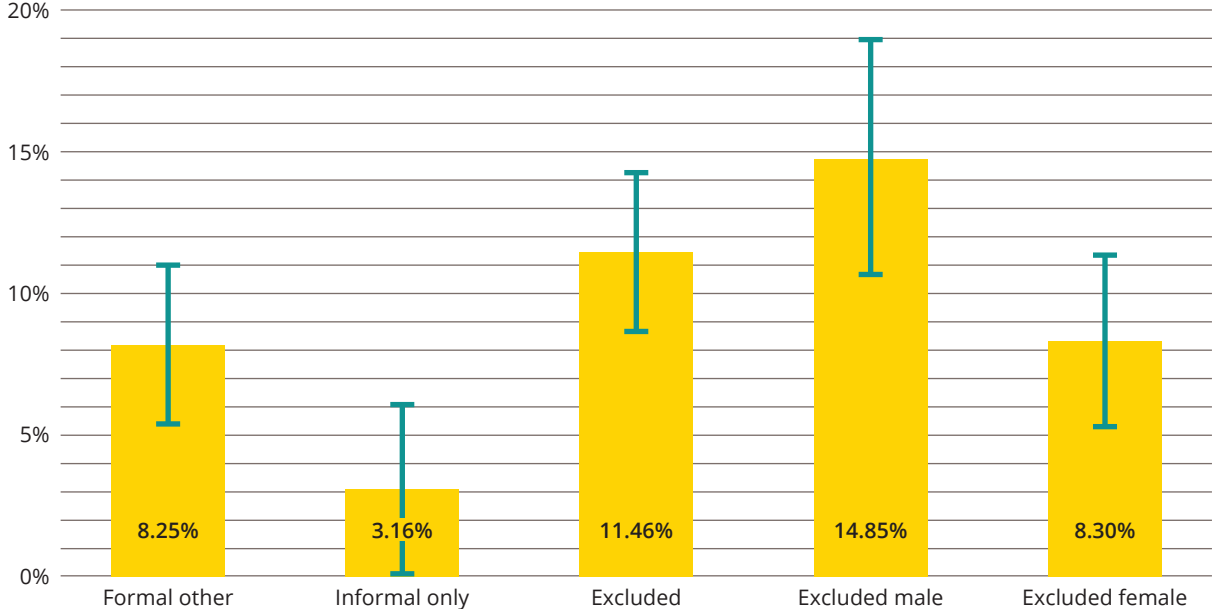
Alternative way to represent the confidence interval

There are many ways to show the interval for key estimates.

South Africa – Access Strand 2016



South Africa – Access Strand 2016



Conclusion

It is important that the users of statistics realise that there is an interval around any estimate (mean or proportion). These intervals can be very small or very large depending on a number of factors. These factors should be identified and considered when interpreting findings. The authors recommend that the interval and MoE be reported on multiple times in a report, especially for key indicators and when the sample size (base) changes due to sub-group analyses.



Bibliography

<http://www.aapor.org/Education-Resources/Election-Polling-Resources/Margin-of-Sampling-Error-Credibility-Interval.aspx>

http://www.scor.qc.ca/en_calculez.html

<http://stattrek.com/statistics/dictionary.aspx?definition=Confidence%20level>

<http://stattrek.com/statistics/dictionary.aspx?definition=Confidence%20level>

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2689604/>

<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4877414/>

<https://www.vox.com/science-and-health/2017/7/31/16021654/p-values-statistical-significance-redefine-0005>

Malhotra, N.K. & Birks, D.F. (2003). Marketing Research. An applied Approach. Prentice Hall.

How to find us:

Get involved. Contact us.

Grant Robertson

E: grant@i2ifacility.org

Bobby Berkowitz

E: bobby@i2ifacility.org

Damola Olowade

E: damola@i2ifacility.org

Shirley Jeffreys-Leach

E: shirley@i2ifacility.org

+27 11 315 9197

i2ifacility.org

 [@i2ifacility](https://twitter.com/i2ifacility)

 [/insight2impact](https://www.linkedin.com/company/i2ifacility)

 [/insight2impact](https://www.facebook.com/i2ifacility)

 [/i2ifacility](https://www.instagram.com/i2ifacility)

Established by



Sponsored by

BILL & MELINDA
GATES *foundation*

