Women's Rights Online
Report Annex

October 2015

This research is part of a 10-country study[1] on Women's Rights Online funded by the Swedish International Development Cooperation Agency (Sida). The study was managed by the World Wide Web Foundation and conducted collaboratively between the Web Foundation, Ipsos MORI and Women's Rights Online partner organisations located in each of the ten countries. The research and analysis took place between January and October in 2015.

[1] Due to delays in fieldwork in Cairo, Egypt was not included in the comparative research report. Please refer to the Egypt country report written by Tadwein Gender Research Centre for data and analysis on Egypt.

ANNEX 1 Country Selection

A total of ten countries across Africa, Asia and Latin America were selected to be included in this phase of the project. While we would have liked to increase the country coverage of this study, we had to limit the study to ten countries due to resource and time constraints.

There are many other countries with equally - if not more - significant needs for data on women's rights online, ICTs and the Internet, and we encourage researchers to use the data from our research to develop a greater database of information on the gender gap in ICT access and use in various countries.

We applied the following criteria to select ten countries for inclusion in this phase of the project.

Criteria for inclusion of countries:

- Countries of programme focus for Sida
- Inclusion of the country in the 2014 Web Index study
- Legal environment that supports equal rights for women and men
- Relatively stable domestic political climate for household research and policy advocacy
- Existing strategic links between World Wide Web Foundation and multi-stakeholder partnerships (E.g Alliance for Affordable Internet)
- A selection of low and middle income countries in various regions (based on World Bank income classification, see table below)
- Availability of secondary data from a variety of sources (World Bank, ITU, WEF Gender Gap index) and national statistical data on ICT/gender
- Partnerships between Web Foundation and local NGOs and research organisations at country level for research implementation and advocacy

Country	**Income Classification**
Mozambique	low income
Uganda	low income
Cameroon	lower-middle income
Egypt	lower-middle income
India	lower-middle income
Indonesia	lower-middle income
Kenya	lower-middle income
Nigeria	lower-middle income
Philippines	lower-middle income
Colombia	upper-middle income

Source: World Bank (http://data.worldbank.org/about/country-and-lending-groups)

ANNEX 2 METHODOLOGY

Co-authored by Daniel Cameron, Sidra Butt-Mughal, and Jennifer Keyes from Ipsos MORI and World Wide Web Foundation staff.

1. Overview of the Survey Methodology

The World Wide Web Foundation held a consultative meeting in December 2014 with Women's Rights Online country partners to discuss the research project and methodologies appropriate to close the gender gap in ICT data. We explored the possibilities of conducting qualitative case study research only, mobile phone based surveys, and surveys focused on both urban and rural areas. Our reasons for choosing surveys in urban areas were twofold. First, high population densities in urban areas made it feasible to base our research on face-to-face interviews, rather than telephone or SMS surveys, leading to more reliable and representative data. Second, because women's choices and experiences around connectivity were a key research focus, we needed to select areas where the basic infrastructure (such as 3G signal coverage and public Internet access points) is available to allow people to connect. This excludes much of the countryside in many developing countries. However, we acknowledge the critical importance of carrying out similar research in rural areas in future.

Ideally, nationally representative surveys, or at least an equally representative sample of rural areas as well as urban areas would have been preferable. However, due to resource and time constraints the surveys were conducted using quota sampling stratified to focus on an urban poor demographic in capital or key cities.

The World Wide Web Foundation and Women's Rights Online country partners designed the questionnaire through a collaborative process involving ongoing consultation from local stakeholders on the phrasing, translation and revision of questions appropriate to city-specific contexts. Country partners conducted pilot surveys prior to the full survey in order to refine further the final questionnaire and strengthen the implementation of the full survey.

Ipsos MORI was responsible for coordinating translation of the survey, designing sampling frameworks, overall management of the data collection and data processing. Data was collected via a quota household face-to-face survey in urban poor areas in cities in 10 countries. The survey was conducted in the capital cities or main economic hubs including: Nairobi, Kampala, Yaoundé, Maputo, Lagos, Greater Cairo, Bogota, Jakarta, New Delhi and Manila.

A quota sample of 1,000 face-to-face interviews was completed with 250 men and 750 women living in poor, urban areas. The focus was on women, hence the larger sample of women and smaller sample of men. Inclusion of a small sample of men is useful for comparative purposes and to provide indication of the gender gap in ICT and Internet use in each city. The data on men was not intended to be representative and should be used with caution.

Given resource constraints, the sample size was 1,000 people in every city. Had we split the sample 50:50, the margin of error would have been relatively large for both groups (male and female). Therefore, given that the project's focus is on women's access to and use of ICTs, and after consultations with in-country partners in each country, it was decided to select the sample to consist of 75% women and 25% men. This was in order to obtain a good representative result of the patterns of access and use of ICTs by women in urban poor areas. Nevertheless, it was also decided to survey a large enough number of men (at least 250 and up to over 300 in some countries)[2] to obtain a good indication of how men access and use ICTs in the same areas, thereby getting some pointers regarding the 'gender gap' in the areas under study.

While the smaller sample of men gives a higher margin of error for every country survey (around 6% for men compared to around 3.6% for women at the 95% confidence level), when we conducted the statistical analysis, this small sample effect no longer held for the global analysis as that was conducted on a sample size of around 2,300 men and 6,800 women.

Critically, even for the individual country samples of over 250 men and 750 women, those appeared to give reliable results, given the relatively homogeneous areas and sections of the population surveyed (urban poor in the capital/main cities). Moreover, we re-calculated the survey results applying a re-weighting to get an indication how the results might change had the sample been 50% women and 50% men (instead of 25% men and 75% women), and the results using the re-weighting were not significantly different in most cases. However, for the final analyses and report, we decided not to use the re-weighted survey results as this method – although statistically valid - has clear limitations from an empirical/observational perspective.

A mixture of Computer Assisted Personal Interviewing (CAPI) and Paper Assisted Personal Interviewing (PAPI) was used to collect the survey data.

[2] The sample size for men in Uganda was 332; Nigeria was 307.

Country	Fieldwork City	Data Collection Method
Cameroon	Yaoundé	CAPI
Colombia	Bogota	PAPI
Egypt	Greater Cairo	PAPI
India	New Delhi	CAPI
Indonesia	Jakarta	PAPI
Kenya	Nairobi	CAPI
Mozambique	Maputo	CAPI
Nigeria	Lagos	CAPI
Philippines	Manila	CAPI
Uganda	Kampala	CAPI

2. Questionnaire

The questionnaire covered topics relating to women and men's access to, use of, and perceived value of the Internet and ICTs. It also looked specifically at the barriers women face in accessing the Internet, where and how women access public information and information on women's rights (including sexual health information and related health services, information on legal rights, gender based violence information and support services). The questionnaire also addressed women's use of the Internet for education, employment and political activity such as voicing opinion and collectively organising on public issues. The questionnaire also looked at women's experiences of harassment via mobile phones and online, uses of social media and perceived value of the web and Internet to daily life. Harassment was defined as including: receiving offensive or threatening calls, texts, emails; having negative, hateful/offensive/ insulting comments posted about you on the Internet; cyberstalking; sexual harassment or luring; non-consensual distribution of photos/videos of you.

“Internet users” were defined as those who have accessed the Internet on any device in the last six months. Using the Internet included all web browsing and communication including Google, Facebook, email, WhatsApp, WeChat, Viber, etc.

A master questionnaire was finalised in English, and was then translated by each Ipsos MORI country field office in the local language (see list below). The translated questionnaires were then checked by a local-language translator and back-checked into English. The World Wide Web Foundation Women’s Rights Online country partners also reviewed, edited, translated and approved the translations of the questionnaire.

Country	Language of Translation
Cameroon	French
Colombia	Spanish
Egypt	Arabic
India	Hindi
Indonesia	Indonesian
Kenya	Swahili
Mozambique	Portuguese
Nigeria	English
Philippines	Tagalog
Uganda	Luganda

The final questionnaires were sent to CAPI scripting teams based in Nairobi, Kenya to be programmed using the Survey To Go / Dooblo application that uses Dimensions software.

The local language links were tested by the Ipsos MORI country research teams to ensure that the questionnaire had been programmed correctly, including logic checks built in to avoid data errors and to minimise the need for data cleaning at the end of fieldwork. The process of testing underwent several rounds of error-checking to ensure that the structure and content of the survey applications for all countries were standard.

3. Ethical Code of Conduct

As a market and social research company, Ipsos MORI abides by the ICC/ESOMAR Code on market and social research. ESOMAR is the world organisation for market and social research, a copy of the guidelines can be found here.

Ipsos MORI's market and social research services meet the requirements of ISO 9001:2008, the international standard for Quality Management Systems; ISO 20252:2006, the international Market Research Standard, which incorporates the Interviewer Quality Control Scheme; ISO 27001:2005, the international standard for Information Security Management Systems.

Ipsos MORI and all fieldwork partners in this project comply with the ESOMAR International Code of Marketing and Social Research Practice. The research was carried out in accordance with the ICC/ESOMAR International Code of Marketing and Social Research Practice.

3.1 Informed Consent

The questionnaire required fieldworkers to obtain verbal informed consent before conducting any interviews. Research participants were provided with a verbal and written overview of the purpose, goals and objectives of the study; assuring they were selected at random. Research participants were assured of their anonymity and privacy in completing the survey. Any participant who declined to give verbal consent to participate in the survey was not interviewed.

4. Pilot Test of the Survey

A pilot phase was conducted between 28th April and 9th May to test the questionnaire in all countries, with the exception of Egypt where the survey was piloted between 10 and 11 June 2015. The pilot in Egypt was slightly delayed due to the time required to obtain government approval of the questionnaire and study as required from the Central Agency for Public Mobilization and Statistics (CAPMAS).

In all cases the pilot was conducted by World Wide Web Foundation's country partners in the respective countries (see list below). Pilot feedback was then used to make revisions to improve the question wording, the translation of certain questions and the flow of the questionnaire.

Country	Fieldwork City	Pilot conducted by Partner Organisation
Cameroon	Yaoundé	I-Vission International
Colombia	Bogota	Fundación Karisma
Egypt	Greater Cairo	Tadwein Gender Research Centre
India	New Delhi	IT for Change
Indonesia	Jakarta	ICT Watch
Kenya	Nairobi	International Association of Women in Radio and Television
Mozambique	Maputo	Science Innovation Information and Communication Technology Research Institute (SIITRI)
Nigeria	Lagos	Paradigm Initiative Nigeria
Philippines	Manila	Foundation for Media Alternatives
Uganda	Kampala	Women of Uganda Network (WOUGNET)

5. Sampling Frameworks

The targeted sample size was 1,000 respondents, including pre-determined quota of 750 women and 250 men. A cross-section of multiple poor urban areas were surveyed in each of the ten cities. The target age group was 18 - 60 years, which included assigned quotas per specific age brackets, largely in line with country demographic patterns where national statistics allowed.

The following data was collected and organised for each country sampling plan:

- Population size of sampling city (urban only) for population ages 18 – 60 years only
- The population size was organised by age and gender (where available)
- List of eligible sampling areas (slum dwelling/ informal settlements)
- Estimated population size of each area (if available for ages 18 -60 years)

- Maps showing outline/ borders of sampling areas and highlighting key landmarks within each sampling area

Population data was collected from publically available sources and preferably from the most recent national census in each country. The data sources for the quotas in each city are listed below for reference.

Population Data Sources for each country

Country	Data Source
Colombia	2013 - Dane Bogotá D.C.: Pobreza Monetaria Comunicado de prensa.
Cameroon	2005 - Source BUCREB: Data from RGPH - Yaoundé Urban Council
Egypt	2006 - Central Agency for Public Mobilization and Statistics (CAPMAS) National Census
India	2011 - National Census - The Registrar General & Census Commissioner, India
Indonesia	2010 - National Census - Central Bureau of Statistics
Kenya	2009 - Census data - Kenya National Bureau of Statistics
Mozambique	2007 - Instituto Nacional De Estatistica – Moçambique, Census Data
Nigeria	2006 - National Bureau of Statistics /National Population Commission (NPC)
Philippines	2010 - National Statistics Office (NSO) - Census Data
Uganda	2014 National Slum Dwellers Federation Of Uganda, Slum Profiles In Kampala

In the following countries, population information was only available at a city-regional level and not at a sub-location level:

1. **Cameroon** – Yaoundé city divided into seven regions; Yaoundé sections one to seven and further into pre-defined clusters.
2. **Egypt** – Cairo city divided into 14 regions and further into sub locations.

3. **India** – Delhi State defined geographically into five districts / regions, Central, North, South, East & West Delhi and then further divided into sub locations.
4. **Indonesia** – Jakarta city defined geographically into five regions, Central, North, South, East & West Jakarta. Each region is divided into districts and then further into villages and neighbourhoods, known as Rukun Tetangga (RT) and Rukun Warga (RW).
5. **Nigeria** – Lagos city divided into 6 regions and a further 9 sub locations. Clusters were created within each sub location.

Quotas for these cities were assigned proportionately to the population size at the lowest available level of geography. These quotas were then divided equally across the sub-locations in each region.

For all other cities, Bogota, Kampala, Manila, Maputo and Nairobi, population data was available for sub-locations and quotas were designed based on this level.

5.1 Sampling stages

Stage 1

- Field offices made a list of **ALL** applicable **urban poor areas** (slum dwelling/ informal settlements) **within the capital city boundaries.**
- All areas listed met the criteria of **urban, poor settlements** where the majority of the population lives below the poverty line.
- Areas that were classified as unsafe or difficult to access were excluded from the listing.
- Where available, each field office provided the estimated population sizes for each sampling area, listed by city region or district and then by sub location.

Stage 2

- Quotas were calculated and assigned proportionate to the applicable population for the sampling regions and sub locations (where available) by age and gender. The quotas were also assigned based on the population profile information available.

Stage 3

- Using Google Maps and/or readily available maps, the boundaries of each sampling area were highlighted
- In countries where maps were unavailable, ground teams surveyed the sampling areas and, with the assistance local administrative leaders, noted landmarks to identify the boundaries of the sampling areas.
- The field supervisor / managers then assigned their team with sampling areas according to the sample and the necessary sample quotas for each area.

- The field supervisor / managers allocated each team starting points within the sampling areas / clusters. These were spread out across the sampling area and assigned to each interviewer.

Stage 5

- In densely populated areas, after every successful interview, interviewers were asked to skip four households and find an interview every fifth house. In cases where they were unsuccessful in securing an interview, they went to the next household interval until they were successful.
- In a household where an interviewer found more than one person who was eligible and willing to participate in the study, the interviewer used the 'birthday rule' and interviewed the man or woman whose birthday had passed most recently.
- The number of interviews achieved around each starting point was never more than 20.
- The details of the sampling region, sub location, starting point / landmark etc. were recorded for every interview.

5.2 Quota sampling

A key of aim of the study was to deliver high quality, quantitative data that is broadly representative of the target population in the ten cities selected for inclusion in the study. Quota sampling is a non-random sampling method and involves a fixed quota of interviews being set (within each sampling point[3]) on variables such as age and gender to ensure the sample is broadly representative of the population of interest. Such quotas are based on the most up-to-date demographic profile of the population in each of the cities. Individual sampling units (households) were randomly selected within the predetermined urban poor areas in each city.

As part of the survey, interviewers asked potential interviewees a series of demographic screening questions to identify whether they fit the profile before continuing with the main survey. Drawing the sample in this way means that it is not possible to calculate selection probabilities, as would be the case in a random probability sampling approach. To reduce interviewer selection bias and increase the randomness of selection, interviewers were given a set of instructions to follow when selecting households, for example, to follow a particular route or skipping a number of households after successful completion of one interview.

[3] "Sample Point" is a clearly defined, specific and unique area that an interviewer will go to achieve a fixed amount of interviews. Each could be a town, village, neighbourhood in a city, street etc.

For this survey, quota sampling was identified as the preferred survey methodology given the difficulty in obtaining a comprehensive sampling frame and because of time and budget constraints.

5.3 Survey Fieldwork

Interviewer training & briefing

Prior to fieldwork in the countries, the Ipsos MORI central project team members held telephone briefings with the project managers in each of country offices in conjunction with the World Wide Web Foundation project team. The telephone briefings were held between the 22nd and 28th April 2015.

The in-country project briefings were held over two to three days and consisted of intensive training sessions, held centrally, usually at the research agency offices. The training sessions in all countries (except Nigeria and Egypt) began on the week commencing 4th May 2015. Training in Nigeria was held in the week commencing 11th May 2015 and training in Egypt started on 10 June 2015.

The trainings were facilitated by the field managers, field coordinators and the World Wide Web Foundation's partners from local organisations in each country. After the questionnaire training session, a half-day pilot was completed. Pilot interviews were reviewed by the field manager, data manager and supervisors, together with the field interviewers, and any issues were raised and discussed during one-on-one sessions with interviewers.

Survey Fieldwork teams

The field team was made up of majority female interviewers due to the importance of women interviewing women and men interviewing men (gender matching criteria). The number of interviewers in each country was between 10 – 25 in total.

To comply with the gender matching interview criteria, female interviewers selected randomly from all female participants in the household, while male interviews selected randomly from all male participants in the household. All sections of the questionnaire were asked to each participant as applicable.

Survey Fieldwork dates

Fieldwork was completed over a four-week period between 11 May and 8 June with the exceptions of Egypt where fieldwork took place between 10 and 28 August. The reason for the delay in Egypt was that after obtaining official permission in June from CAPMAS to

conduct the survey, the fieldwork had to be further postponed until after Ramadan and Eid holidays.

Survey Interview duration

The English version of the questionnaire was tested and took approximately 25-30 minutes to administer. Translated versions of the questionnaires ran for 30-35 minutes on average.

Substitution and replacement of sampling units

Field agencies were not authorised to substitute or replace any of the regions or sub locations from those listed and approved in the sampling plans. Any requests for substitution were submitted to the project manager and then assessed for replacement.

Survey Fieldwork challenges and limitations

Despite fieldwork being successful, the following challenges were encountered during the data collection process:

- **Safety and Security** – Several interviewers expressed concerns about their personal security while conducting fieldwork in the slum areas. There were several incidences where participants were harsh and unwelcoming towards the interviewers. This may be largely attributed to concerns about safety in these areas where insecurity is high and 'outsiders' are viewed with suspicion.
- **Environment –** In Kenya and Uganda, fieldwork was affected by heavy rains which made access to the slum areas difficult. In New Delhi, the fieldwork period was disrupted due to a heat wave which made travel and fieldwork extremely difficult. In Nigeria, a severe fuel shortage hindered and delayed fieldwork by one week.
- **Refusals –** There were cases where participants, both potential and actual, demanded incentives to participate (no incentives were offered anywhere). Most participants, however, agreed to continue without incentive and there were only a few cases where participants refused to be interviewed on the basis that incentives were not being offered.
- **Fieldwork in Egypt -** The Egyptian Government requires study permissions to conduct fieldwork in Egypt. The process of obtaining permissions led to initial delays in the study timeline for Egypt. We were also requested to remove or change (rephrase) the following parts of the questionnaire to obtain approval which influenced the comparability of these survey questions across countries:
    - The title of the survey which was "Survey on Women's Empowerment through the Internet" (changed to "Survey on using the Internet")

- Option to select “Military Personnel” in the question on “type of employment” (removed)
- Survey question on the personal value of the Internet to practice religion (removed)
- Survey question on where people have looked for information on drug and alcohol abuse (removed)
- Question on where people have gone to report corruption (removed)
- Question on people’s opinion about government regulation of media (removed)[4] and on media freedom (rephrased from “The media should have the right to publish views and ideas without government interference to “The media should have the right to publish views and ideas”)
- Reference to term “sexual” in terms of sexual harassment (re-translated as colloquial term for “harassment” which is commonly understood to include sexual harassment)

6. Quality control in Survey Research

6.1 Field checks

Quality control measures were implemented and included telephone call-backs, in addition to physical back-checks to some participants to validate some of their survey responses. Both telephone and physical checks were randomised, without any prior interviewer notification. Participants were asked for the gender of the interviewer that visited them, the length of the interview, the topic of the survey as well as a few main questions from the questionnaire. Participant screening information was also sometimes requested. On average, between 10 and 30% of participants were contacted through telephone call backs, while a minimum of 5% of participants were reached through physical back checks.

Country field offices were required to do a minimum of 10% back checks, however a large number of participants were often unreachable for telephonic back-checks. This was due to poor network coverage or a working telephone number not being available, among other reasons. To address this, telephonic back checks were often supplemented by supervisor-accompanied visits and return visits to the field.

[4] Survey question Q25-D “The government should have the right to prevent the media from publishing things that it considers harmful to society.”

6.2 Data & validation checks

During fieldwork, field supervisors often downloaded and checked data to ensure that any errors or queries could be flagged early. Data was received in real time as the CAPI devices were set to synchronize automatically. The data was then downloaded in the country offices and checked for completeness and validity.

The following checks were also implemented to ensure quality:

- **GPS location services** – this was recorded using the CAPI devices in some countries. Back checkers were able to download the collected data and verify that the interviewers are indeed at the required sampling areas, as well as trace the routes they took.
- **Interview length** – the data collection devices record the start and stop of every questionnaire, checking on questionnaire length. This is an important part of quality checks to identify outliers and surveys that would need to be further validated.

The interviews/cases found not to comply with the quality checks were removed from the data; some of the issues were supported by telephonic and field back checks.

6.3 Data processing

Once fieldwork checks were completed, the data was retrieved and downloaded for all countries and reformatted to ensure that the data structure for all countries was the same. This was to ensure comparability across all countries.

Ipsos MORI's data processing team then worked on labelling and processing the data before compiling data tables with the key variables across all countries. The variables for the data tables included:

- gender
- age
- Internet use
- education
- marital status
- employment status
- personal mobile phone ownership
- mobile phone access,
- income/poverty level (includes: primary income earner, electricity available, cooking fuel type, ownership of mode of transport, type of floor in house, type of roof in the house)

Data tables were constructed in Excel format. The full merged data set containing the data for all countries was structured in SPSS format for additional analysis by the World Wide Web Foundation statistics team.

7. Qualitative Research Methodology

In addition to our 10-country survey research, Women's Rights Online country partner organisations designed and conducted qualitative research which included key informant interviews, focus group discussions, and issue-based case studies.

Among other purposes, the qualitative research was intended to:
a) compliment the survey data through narrative and ethnographic techniques,
b) delve deeper into specific topics identified as priorities by country partners,
c) lend greater personal insight into the priority policy issues for women's rights online.

Please refer to the individual country reports for details on the qualitative research methodologies employed by each project implementing partner.

8. Some Basic Assumptions

- Population data used in sampling frameworks is correct as per each country's most recent population census
- Urban poor areas are correctly identified based on local census and local knowledge

9. Limitations to the study sample and design

- The study is not fully representative of the male or female populations at national level:
    - The study excluded rural populations
    - The study excluded middle and higher income groups
    - The sample size of males interviewed was relatively small compared to the sample size of women interviewed, which means, among other things, that the margin of error for the results for the males interviewed is larger than that for the females.
- For several survey questions the options generated small numbers of respondents. We were cautious in the analysis of responses that generated less than 50

respondents in total, and considered any sample less than 50 to be too small to draw clear conclusions.

10. Statistical analysis - Methods used for data analysis

Univariate, bivariate and logistic regressions were used for data analysis. In univariate analysis, we used bar-graphs and for bivariate analysis, cross tabulation were used to describe the relationship between the dependent variables and the independent variables.

In multivariate analysis, we used logistic regression models and computed odds ratio OR (controlled effect) with 95% confidence interval for the association between dependent variable and each independent variable. This technique was used because the dependent variables were binary (Yes / No; user/non-users). Results of multivariate logistic regressions provide the effects of each of the predictors in the dependant variable when the effects of other variables are controlled.

An odds ratio (OR) is a measure of association between a factor and an outcome. The OR represents the odds that an outcome will occur given a particular factor, compared to the odds of the outcome occurring in the absence of that factor. Odds ratios are most commonly used in case-control studies, however they can also be used in cross-sectional and cohort study designs as well (with some modifications and/or assumptions). The odds ratio can also be used to determine whether a particular predictor is a risk factor for a particular outcome, and to compare the magnitude of various risk factors for that outcome.

- OR=1 predictor does not affect odds of outcome
- OR>1 Predictor associated with higher odds of outcome
- OR<1 Predictor associated with lower odds of outcome

11. Note on construction of Wealth Index

The survey included questions designed to assess household income/wealth. Although the focus areas were the urban poor in the main/capital cities, there are variations in income and wealth levels within this group. This was an important variable to assess, as there is a strong relationship between economic status and Internet use/access. We wanted to assess whether women's access and use of the Internet is hampered by factors specifically related to gender, as well as income and education-related factors.

It is widely acknowledged (see for example, DHS and Unicef work on the subject) that data collected on income (i.e. self-reporting) in household surveys in developing countries (and elsewhere) show substantial measurement errors and reporting bias when people are asked directly about their incomes in surveys. Asset-based approaches (as proxies) have been shown to be relatively more accurate in assessing household living standards and economic status (and hence poverty levels) in country-specific surveys. These could be used to construct country-specific or global wealth indexes. The latter have some shortcomings (mainly the loss of country-specific information), but are useful in many instances for cross-country comparisons (see, for example, the International Wealth Index), especially when the analysis is combined with country-specific indexes.

For the surveys we conducted and the analysis of the results, we included a few questions on some durable household assets and household characteristics, albeit not a comprehensive set of questions given resource constraints. Subsequently, we constructed both a country-specific wealth index for every country, as well as a global index for the nine countries under study, which were used in the testing and analysis of the results.

We collected data for income/household wealth proxies including information on employment status, as well as the following household assets: type of transport owned by the household, type of roof, type of floor and type of cooking fuel used by the household. Based on feedback from our country partners on the local context and status of each household asset, we allocated each of the assets into the different income categories (low, middle and high), assigning grade 1 for poor, 2 for middle and 3 for high, in line with which category of the population class usually owns that asset. Therefore, owning an asset will provide the grade of 1, 2 or 3, and not owning an asset will not provide any grade (marked as NA or missing).

The next step consisted of taking the average of the grades assigned to each individual to determine the wealth index for each individual. Example: A person owing a car (3), tile roof (3) and cooks with gas (2) and answered don’t know about her floor (NA), will have a grade of (3+3+2)/3=2.7

We created the indicator 'Wealth Index' for each country based on the wealth proxies above. We then grouped the information (based on common gradings across countries for poor, middle and high) to create a single global wealth index variable (called “Wthcountryglobgrp” in the regression analysis). The single global wealth index variable takes all the lower economic status individuals in each and every country, all the middle economic status individuals of every country, and all the higher economic status individuals for every country.

ANNEX 3: Results on Gender Gap Analysis and Key Variables

We tested many hypotheses using the results of the survey, but of most relevance here are the tests and results relating to gender discrimination in Internet use and access.

Overall, the results indicate that, controlling for the effects of education and income (both as indicated by employment status and the wealth level of households), women are on average 47% less likely to use the Internet than men.

(See equations below - the odds ratios.)

This increases to 51% when also controlling for age, but we decided to omit age as it is a biological given rather than a vector of socially determined inequality.

.....

Regression results

Main Variables:

Q12 is the survey question relating to Internet use “Have you accessed the Internet using any device in the past six months?” Please note: using the Internet includes Facebook, email, Whatsapp, WeChat, Viber, Google search, etc. or their local equivalents.

Wthcountryglobgrp: is the household wealth indicator

Primary, secondary, tertiary: are education levels

logistic q12_an i.s2 i.q1 i.q2_an4 i.Wthcountryglobgrp i.workstatg

```
logistic q12_an i.s2 i.q1 i.q2_an4 i.Wthcountryglobgrp i.workstatg

Logistic regression                             Number of obs   =       9623
                                                LR chi2(16)     =    2806.21
                                                Prob > chi2     =     0.0000
Log likelihood = -5148.2444                     Pseudo R2       =     0.2142

------------------------------------------------------------------------------
             q12_an | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
```

```
-----------------------------+------------------------------------------------------------------
                          s2 |
                      Female |   .4925599   .0288165   -12.10   0.000     .4391985    .5524046
                             |
                          q1 |
                18 - 24 years |   6.253669   1.000111    11.46   0.000      4.57096    8.555835
                25 - 29 years |   4.369163   .7063205     9.12   0.000     3.182678    5.997963
                30 - 34 years |   3.331441    .547339     7.32   0.000     2.414259     4.59706
                35 - 39 years |   2.727468   .4585228     5.97   0.000     1.961831    3.791907
                40 - 44 years |   1.969521   .3427685     3.89   0.000     1.400298    2.770133
                45 - 49 years |   1.648348   .2997009     2.75   0.006     1.154209    2.354038
                50 - 54 years |   1.291751     .24753     1.34   0.182     .8872963    1.880569
                             |
                      q2_an4 |
                     Primary |   4.253597   1.799247     3.42   0.001     1.856535    9.745624
                   Secondary |   18.27707    7.6699     6.92   0.000     8.029775    41.60159
                    Tertiary |   76.50878   32.47776    10.22   0.000     33.29497    175.8101
                             |
             Wthcountryglobgrp |
                           2 |   1.450779   .1204089     4.48   0.000     1.232976    1.707055
                           3 |   1.571797   .1403546     5.06   0.000     1.319435    1.872428
                             |
                    workstatg |
Wage employed or self emp..  |   .8695923   .0656536    -1.85   0.064     .7499816    1.008279
                     Student |   1.413687   .1564173     3.13   0.002     1.138078     1.75604
              Women Homemaker |   .8714915   .0743048    -1.61   0.107     .7373748    1.030002
                             |
                       _cons |   .0170417   .0077337    -8.97   0.000     .0070021     .041476
-------------------------------------------------------------------------------------------------
```

logistic q12_an i.s2 i.q2_an4 i.WthcountryglobgrpF

```
logistic q12_an i.s2 i.q2_an4 i.WthcountryglobgrpF

Logistic regression                             Number of obs   =       9623
                                                LR chi2(5)      =    2213.43
                                                Prob > chi2     =     0.0000
Log likelihood = -5444.6377                     Pseudo R2       =     0.1689

------------------------------------------------------------------------------
             q12_an | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
--------------------+---------------------------------------------------------
                 s2 |
             Female |   .5355381   .0283025   -11.82   0.000     .4828424    .5939847
                    |
             q2_an4 |
            Primary |   4.368471   1.832239     3.52   0.000     1.920063    9.939015
          Secondary |   24.16439   10.05637     7.65   0.000     10.68897    54.62808
           Tertiary |   104.1883   43.83574    11.04   0.000     45.67599    237.6566
                    |
2.WthcountryglobgrpF |   1.413585   .1119502     4.37   0.000     1.210348    1.650949
               _cons |   .0424553   .0179238    -7.48   0.000     .0185597    .0971164
------------------------------------------------------------------------------
```

logistic q12_an i.s2 i.q2_an4 i.Wthcountryglobgrp

```
logistic q12_an i.s2 i.q2_an4 i.Wthcountryglobgrp

Logistic regression                             Number of obs   =       9623
```

```
                                                LR chi2(6)      =    2214.99
                                                Prob > chi2     =     0.0000
Log likelihood = -5443.8566                     Pseudo R2       =     0.1690

------------------------------------------------------------------------------
          q12_an | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----------------+------------------------------------------------------------
              s2 |
          Female |   .5342501   .0282542   -11.85   0.000     .4816461    .5925992
                 |
          q2_an4 |
         Primary |   4.367588   1.831881     3.51   0.000     1.919664    9.937061
       Secondary |   24.13943   10.04604     7.65   0.000     10.67788    54.57192
        Tertiary |   103.8286   43.68534    11.03   0.000     45.51754    236.8403
                 |
Wthcountryglobgrp |
               2 |   1.384405   .1120598     4.02   0.000     1.181309     1.62242
               3 |   1.477763   .1282233     4.50   0.000     1.246658    1.751711
                 |
           _cons |    .042572   .0179733    -7.48   0.000     .0186105    .0973846
------------------------------------------------------------------------------
```

logistic q12_an i.s2
logistic q12_an i.s2

```
Logistic regression                     Number of obs   =       9640
                                        LR chi2(1)      =     251.10
                                        Prob > chi2     =     0.0000
Log likelihood = -6435.4183             Pseudo R2       =     0.0191

------------------------------------------------------------------------------
      q12_an | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          s2 |
      Female |   .4732837   .0224597   -15.76   0.000     .4312486     .519416
       _cons |    1.26389   .0514719     5.75   0.000     1.166928    1.368909
------------------------------------------------------------------------------
```

logistic q12_an i.q2_an4
logistic q12_an i.q2_an4

```
Logistic regression                     Number of obs   =       9623
                                        LR chi2(3)      =    2052.24
                                        Prob > chi2     =     0.0000
Log likelihood =  -5525.233             Pseudo R2       =     0.1566

------------------------------------------------------------------------------
      q12_an | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      q2_an4 |
     Primary |    4.82107   2.019286     3.76   0.000     2.121392    10.95635
   Secondary |   28.01863   11.64227     8.02   0.000     12.40957    63.26115
    Tertiary |    126.892   53.29113    11.53   0.000      55.7126    289.0116
             |
       _cons |    .031579    .013094    -8.33   0.000     .0140104    .0711776
------------------------------------------------------------------------------
```

logistic q12_an i.Wthcountryglobgrp
logistic q12_an i.Wthcountryglobgrp

```
Logistic regression                             Number of obs   =       9640
                                                LR chi2(2)      =     157.64
                                                Prob > chi2     =     0.0000
Log likelihood =  -6482.145                     Pseudo R2       =     0.0120

------------------------------------------------------------------------------
      q12_an | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
Wthcountryglobgrp |
           2 |   2.210042   .1587179    11.04   0.000     1.919863    2.544082
           3 |   2.461383   .1895366    11.70   0.000     2.116573    2.862367
             |
       _cons |   .3482759   .0232323   -15.81   0.000     .3055925     .396921
------------------------------------------------------------------------------
```

logistic q12_an i.workstatg

```
logistic q12_an i.workstatg

Logistic regression                             Number of obs   =       9640
                                                LR chi2(3)      =     483.91
                                                Prob > chi2     =     0.0000
Log likelihood = -6319.0136                     Pseudo R2       =     0.0369

------------------------------------------------------------------------------
                   q12_an | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
--------------------------+---------------------------------------------------------------
                workstatg |
Wage employed or self e~d |   .7425914   .0472423    -4.68   0.000     .6555382     .841205
                  Student |   3.361963   .3316142    12.29   0.000     2.770977    4.078993
          Women Homemaker |   .5352784   .0388549    -8.61   0.000     .4642934    .6171162
                          |
                    _cons |   .8952234   .0511297    -1.94   0.053     .8004164     1.00126
------------------------------------------------------------------------------
```

logistic q12_an i.s2 i.q1

```
logistic q12_an i.s2  i.q1

Logistic regression                             Number of obs   =       9640
                                                LR chi2(8)      =    1235.36
                                                Prob > chi2     =     0.0000
Log likelihood = -5943.2877                     Pseudo R2       =     0.0941

------------------------------------------------------------------------------
        q12_an | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
---------------+--------------------------------------------------------------
            s2 |
        Female |   .4301258   .0217832   -16.66   0.000     .3894821    .4750109
               |
            q1 |
18 - 24 years  |   11.28306   1.658303    16.49   0.000     8.459083    15.04979
25 - 29 years  |    7.26833   1.084121    13.30   0.000     5.425903    9.736373
30 - 34 years  |   4.977658   .7540356    10.59   0.000     3.698977     6.69836
35 - 39 years  |   3.851654   .5966576     8.71   0.000     2.843074    5.218028
```

```
40 - 44 years  |   2.622067   .4211332     6.00   0.000     1.913953    3.592164
45 - 49 years  |   2.109353   .3539829     4.45   0.000     1.518113    2.930855
50 - 54 years  |   1.650638   .2907835     2.84   0.004     1.168693    2.331327
               |
         _cons |   .2589539   .0373041    -9.38   0.000     .1952545    .3434345
------------------------------------------------------------------------------
```

logistic q12_an i.s2 i.q2_an4

```
logistic q12_an i.s2  i.q2_an4

Logistic regression                               Number of obs   =       9623
                                                  LR chi2(4)      =    2193.89
                                                  Prob > chi2     =     0.0000
Log likelihood = -5454.4048                       Pseudo R2       =     0.1674

------------------------------------------------------------------------------
      q12_an | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          s2 |
      Female |   .5348131    .028245   -11.85   0.000     .4822227    .5931388
             |
      q2_an4 |
     Primary |    4.43704   1.860199     3.55   0.000      1.95089    10.09146
   Secondary |   25.51389   10.61057     7.79   0.000     11.29234    57.64605
    Tertiary |   111.2092   46.75157    11.21   0.000     48.78671    253.5011
             |
       _cons |    .055214   .0230483    -6.94   0.000     .0243628    .1251328
------------------------------------------------------------------------------
```

logistic q12_an i.s2 i.Wthcountryglobgrp

```
logistic q12_an i.s2  i.Wthcountryglobgrp

Logistic regression                               Number of obs   =       9640
                                                  LR chi2(3)      =     401.26
                                                  Prob > chi2     =     0.0000
Log likelihood = -6360.3369                       Pseudo R2       =     0.0306

------------------------------------------------------------------------------
           q12_an | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
------------------+----------------------------------------------------------------
               s2 |
           Female |   .4757432   .0227665   -15.52   0.000     .4331502    .5225244
                  |
Wthcountryglobgrp |
                2 |   2.156183   .1565512    10.58   0.000     1.870181    2.485924
                3 |   2.459701   .1915308    11.56   0.000      2.11155    2.865255
                  |
            _cons |   .6122924   .0465347    -6.45   0.000      .527554    .7106419
------------------------------------------------------------------------------
```

logistic q12_an i.s2 i.workstatg

```
logistic q12_an i.s2  i.workstatg

Logistic regression                               Number of obs   =       9640
```

```
                                                LR chi2(4)      =     681.13
                                                Prob > chi2     =     0.0000
Log likelihood = -6220.4003                     Pseudo R2       =     0.0519

------------------------------------------------------------------------------
             q12_an | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
--------------------+---------------------------------------------------------
                 s2 |
             Female |   .4875059   .0251134   -13.95   0.000     .4406879    .5392978
                    |
          workstatg |
Wage employed or s.. |   .6967625   .0450636    -5.59   0.000     .6138082    .7909279
            Student |   3.373298   .3361626    12.20   0.000     2.774783    4.100911
     Women Homemaker |   .6498221   .0483744    -5.79   0.000     .5616023       .7519
                    |
              _cons |   1.508999   .1041335     5.96   0.000     1.318102    1.727543
------------------------------------------------------------------------------
```

logistic q12_an i.s2 i.q1 i.q2_an4 i.WthcountryglobgrpF i.workstatg

```
logistic q12_an i.s2 i.q1 i.q2_an4 i.WthcountryglobgrpF i.workstatg

Logistic regression                             Number of obs   =       9623
                                                LR chi2(15)     =    2804.04
                                                Prob > chi2     =     0.0000
Log likelihood = -5149.3317                     Pseudo R2       =     0.2140

------------------------------------------------------------------------------
             q12_an | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
--------------------+---------------------------------------------------------
                 s2 |
             Female |    .494112   .0288819   -12.06   0.000     .4406267    .5540896
                    |
                 q1 |
      18 - 24 years |   6.234023   .9969888    11.44   0.000     4.556573    8.529008
      25 - 29 years |   4.363552   .7054946     9.11   0.000     3.178475    5.990479
      30 - 34 years |   3.328341   .5468709     7.32   0.000     2.411955    4.592895
      35 - 39 years |   2.727474   .4585614     5.97   0.000     1.961783    3.792018
      40 - 44 years |   1.972958   .3433962     3.90   0.000       1.4027    2.775049
      45 - 49 years |   1.646326   .2993521     2.74   0.006     1.152767    2.351203
      50 - 54 years |   1.290288   .2473013     1.33   0.184     .8862214    1.878586
                    |
             q2_an4 |
            Primary |   4.258799   1.801357     3.43   0.001     1.858882    9.757139
          Secondary |   18.31476   7.685362     6.93   0.000     8.046634     41.6858
           Tertiary |   76.83981   32.61644    10.23   0.000      33.4406    176.5625
                    |
2.WthcountryglobgrpF |   1.487853   .1208069     4.89   0.000     1.268955    1.744511
                    |
          workstatg |
Wage employed or s.. |   .8649482   .0652112    -1.92   0.054     .7461314    1.002686
            Student |   1.411578   .1561366     3.12   0.002     1.136455    1.753305
     Women Homemaker |   .8666448   .0738061    -1.68   0.093     .7334158    1.024076
                    |
              _cons |   .0170705    .007747    -8.97   0.000     .0070137    .0415471
------------------------------------------------------------------------------
```

logistic q12_an i.s2 i.q1 i.q2_an4 i.WthcountryglobgrpF

```
Logistic regression                             Number of obs   =       9623
                                                LR chi2(12)     =    2775.96
                                                Prob > chi2     =     0.0000
```

```
Log likelihood = -5163.3687                     Pseudo R2       =     0.2119

------------------------------------------------------------------------------
             q12_an | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
--------------------+---------------------------------------------------------
                 s2 |
             Female |   .4937464   .0271314   -12.84   0.000     .4433333    .5498922
                    |
                 q1 |
      18 - 24 years |   6.853398   1.089668    12.11   0.000     5.018428    9.359318
      25 - 29 years |   4.383312   .7086978     9.14   0.000     3.192856    6.017629
      30 - 34 years |    3.26498   .5360773     7.21   0.000     2.366583    4.504425
      35 - 39 years |   2.670194   .4485864     5.85   0.000     1.921069    3.711442
      40 - 44 years |   1.912095   .3323996     3.73   0.000     1.359991    2.688331
      45 - 49 years |     1.6014   .2911201     2.59   0.010     1.121396    2.286866
      50 - 54 years |   1.266422    .243221     1.23   0.219     .8691648    1.845248
                    |
             q2_an4 |
            Primary |    4.21949   1.784786     3.40   0.001     1.841677    9.667328
          Secondary |   18.62456    7.81478     6.97   0.000     8.183246    42.38833
           Tertiary |   81.15193   34.43435    10.36   0.000     35.32787    186.4148
                    |
2.WthcountryglobgrpF |   1.522819   .1236946     5.18   0.000     1.298696    1.785621
               _cons |   .0148182   .0066827    -9.34   0.000     .0061224    .0358649
------------------------------------------------------------------------------
```

logistic q12_an i.s2 i.q1 i.q2_an4 i.Wthcountryglobgrp

```
logistic q12_an i.s2 i.q1 i.q2_an4 i.Wthcountryglobgrp

Logistic regression                             Number of obs   =       9623
                                                LR chi2(13)     =    2778.78
                                                Prob > chi2     =     0.0000
Log likelihood = -5161.9633                     Pseudo R2       =     0.2121

------------------------------------------------------------------------------
             q12_an | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
--------------------+---------------------------------------------------------
                 s2 |
             Female |   .4920362   .0270626   -12.89   0.000     .4417534    .5480424
                    |
                 q1 |
      18 - 24 years |   6.873117   1.092728    12.12   0.000     5.032976    9.386046
      25 - 29 years |   4.391554   .7099402     9.15   0.000     3.198989    6.028702
      30 - 34 years |   3.271229    .537061     7.22   0.000     2.371172    4.512932
      35 - 39 years |   2.672484   .4489232     5.85   0.000     1.922784    3.714494
      40 - 44 years |    1.91057   .3320913     3.72   0.000     1.358967    2.686068
      45 - 49 years |     1.6052   .2917874     2.60   0.009     1.124089    2.292226
      50 - 54 years |   1.268825   .2436185     1.24   0.215     .8708998    1.848567
                    |
             q2_an4 |
            Primary |   4.214578   1.782789     3.40   0.001     1.839464    9.656436
          Secondary |   18.58201   7.797267     6.96   0.000     8.164259    42.29301
           Tertiary |   80.69755   34.24385    10.35   0.000      35.1281    185.3814
                    |
Wthcountryglobgrp |
                  2 |   1.479549    .122866     4.72   0.000     1.257312    1.741067
                  3 |    1.62002   .1444974     5.41   0.000     1.360184    1.929493
                    |
              _cons |   .0148612   .0067018    -9.33   0.000     .0061404    .0359674
```

ANNEX 4: OVERVIEW OF DEMOGRAPHIC DATA

Education level of Survey Sample

Education Level	% Men Sampled	% Women Sampled
Some primary level education	8	12
Finished primary level education	12	16
Some secondary education	24	28
Finished secondary level education	34	29
Some tertiary education	14	9
Completed tertiary education	6	3
No formal education	1	2

% Men and Women surveyed who currently earn an income

	% Men	% Women
Yaounde, Cameroon	61	42
Bogotá, Colombia	62	43
New Delhi, India	68	36
Jakarta, Indonesia	91	55
Nairobi, Kenya	89	52
Maputo, Mozambique	53	46
Lagos, Nigeria	66	68
Manila, Philippines	67	35
Kampala, Uganda	84	58
TOTAL %	72	49

Working Status of the Survey Sample

	% Men sampled	% Women sampled

Wage employed (full time / part time)	38	17
Self employed	38	31
Student (Full time / part time)	10	8
Homemaker / housewife / husband	0	32
Unemployed / not currently working (in between jobs)	11	11
Other	3	1