Butler, Caroline

Conference Paper

Disaggregating the drivers of mobile technology adoption: the threat of unobservable gender biases

ITS Online Event, 14-17 June 2020

Provided in Cooperation with:
International Telecommunications Society (ITS)

Suggested Citation: Butler, Caroline (2020) : Disaggregating the drivers of mobile technology adoption: the threat of unobservable gender biases, ITS Online Event, 14-17 June 2020, International Telecommunications Society (ITS), Calgary

This Version is available at:
https://hdl.handle.net/10419/224848

Caroline Butler, GSMA Intelligence

Working Paper

Disaggregating the drivers of mobile technology adoption: the threat of unobservable gender biases

Abstract

As the reach of mobile technology grows, it is becoming an increasingly powerful tool for access to welfare-enhancing information and services in low- and middle-income countries (LMICs). However, digital inclusion remains far from universal. Across LMICs, 14 per cent of adults still do not own a mobile phone, 39 per cent do not use mobile internet, and 38 per cent do not own a smartphone. Among other characteristics, these digitally excluded individuals are predominantly female.

This study seeks to better understand the key drivers of mobile ownership, mobile internet, and smartphone use, with a particular focus on gender. Discrete-choice models, including logit, probit and linear models, are used to estimate the probability of adoption of these three types of technology. By including a suite of control variables for observable drivers of mobile adoption (e.g. education levels, age, employment, rural-urban location), the coefficient for gender represents non-observable effects which could be a product of discrimination and cultural norms. Furthermore, importance is placed on the inclusion of interaction terms in the regressions (for example, gender interacted with rural location), in order to isolate different degrees of marginalisation across the female population.

In addition to the focus on gender, the marginal effects of the dependent variables for other factors (such as geography, education, employment, and age) will aid understanding of the key predictors of mobile use more generally. This research also shows how these predictors might vary by country and region, how they relate to each other, and which are the most important. This will provide relevant and important information for policy-makers.

The research makes use of multiple years (2017, 2018, and 2019) of data from face-to-face consumer surveys sourced from the GSMA, which includes nationally representative samples of at least 1,000 respondents for 31 low- and middle-income countries. The wide geographic scope, and multi-year nature of the survey data results in a unique contribution to the literature, and the substantial number of observations allows for novel analysis of intersections of the female population.

In summary, the initial results find that: women are 5 percentage points (pp) less likely to own a mobile phone then men, 6pp less likely to use mobile internet, and 4pp less likely to own a smartphone, even when other relevant socioeconomic and demographic factors are controlled for. This unobservable gender effect is more pronounced in certain regions, especially South Asia, but with no significant link in Latin America and Caribbean. The marginal effects of the interaction variables indicate that the negative impact is enhanced for women that live in rural areas, have low levels of literacy, and are not working.

In addition, this study finds that the probability of mobile technology adoption increases (with varying magnitudes by technology type and region) with income, education, urban location, literacy, and employment. Adoption of mobile technology largely declines with age, but the impact generally does not appear to start until age 45 and above for mobile ownership.

Contents

1 Introduction and key findings

Recent growth in the mobile industry has driven an unprecedented increase in digital inclusion in low- and middle-income countries (LMICs) (GSMA, 2019). At the end of 2019, 3.8 billion adults owned a mobile phone in LMICs, 2.7 billion used mobile internet, and 2.7 billion owned a smartphone. As the reach of mobile grows, it is becoming an increasingly powerful tool for delivering life-enhancing information, services and opportunities to millions who have not had the opportunity to access them before.

However, despite this growth, mobile ownership remains far from universal. Across LMICs, 14 per cent of adults still do not own a mobile phone, 39 per cent do not use mobile internet and 38 per cent do not own a smartphone.[1] These individuals tend to belong to the most marginalised groups: they are disproportionately rural, illiterate and older, and they are also predominantly female.

This research seeks to better understand the key drivers of mobile ownership, mobile internet, and smartphone use, with a particular focus on gender. By using a suite of control variables for observable drivers of mobile adoption (e.g. education levels, age, employment, rural-urban location), the coefficient for gender, which is the independent variable of interest, should represent if there are non-observable aspects related to gender, and the magnitude of these effects.[2] Previous studies have found mixed results on this topic. For example, studies in Nigeria, Senegal, Burkina Faso, and Tanzania found that gender was associated with lower likelihood of owning a mobile phone.[3] However, there was no significant gender effect in studies in Gabon, South Africa and Mozambique.[4] The majority of studies have thus far focused on a single country and survey. A summary of relevant literature can be found in Annex 2.

The econometric analysis makes use of multiple years (2017, 2018, and 2019) of GSMA Intelligence consumer survey data, which covers 31 LMICs (see Annex 1 for a list of countries). This analysis will both inform the robustness, and build on previous studies, by utilising this dataset that includes both a wide range of LMICs over three years, with detailed survey questions on mobile technology adoption. Furthermore, this analysis could be further augmented by utilising this survey data that covers types of usage of mobile technology, and reported barriers to adoption of the technologies. Aside from the gender focus, the outputs from this analysis will aid understanding of the key predictors of mobile use more generally (and how these might vary by region), how they relate to each other, and which are the most important. This will provide important information for policy purposes.

In summary, we find that:

- Women are less likely to use mobile, mobile internet and smartphones than men, even when other relevant socioeconomic and demographic factors such as income, education and geography are controlled for. Women are 5 percentage points (pp) less likely to own a phone then men (other factors being equal), 6pp less likely to use mobile internet, and 4pp less likely to own a smartphone.
- This unobservable gender effect is more pronounced in certain regions, especially South Asia, but with no significant link in Latin America and Caribbean.
- The negative effect is enhanced for women that live in rural areas, have low levels of literacy and are not working.

[1] GSMA Intelligence modelling based on the GSMA Intelligence consumer survey.
[2] For example, potentially based on cultural norms and discrimination.
[3] See Forenbacher et al., 2019; Gillwald et al., 2010; Chabossou et al., 2009
[4] See Penard et al., 2012; Gillwald et al., 2010; Chabossou et al., 2009

2 Data and methodology

This study primarily makes use of the GSMA Intelligence Global Consumer Survey (2017, 2018 and 2019), carried out by Ipsos MORI. In all countries, a nationally representative sample of the adult population[5] was selected. At least 1,000 interviews were conducted in each country surveyed, with 2,000 conducted in India and China. To achieve a nationally representative sample, quotas were applied in line with census data on the following metrics: age category by gender; urban and rural distribution by gender; region/state, and socio-economic class (SEC) to ensure a representative portion of lower income respondents were included.

The surveys use a mix of purposive and random sampling approaches. Depending on the country, sampling points were either randomly distributed with an administrative area's probability of selection proportionate to the size of its population (random sampling), or selected to reflect the linguistic, cultural and economic variations of each country (purposive sampling). Based on the selection of sampling points above we can do similar and stratify the data to urban/rural areas.[6] Within sampling points, systematic random routes were used for residence selection.[7]

Chart 1: Summary of technology adoption by gender in the GSMA Intelligence Global Consumer Survey

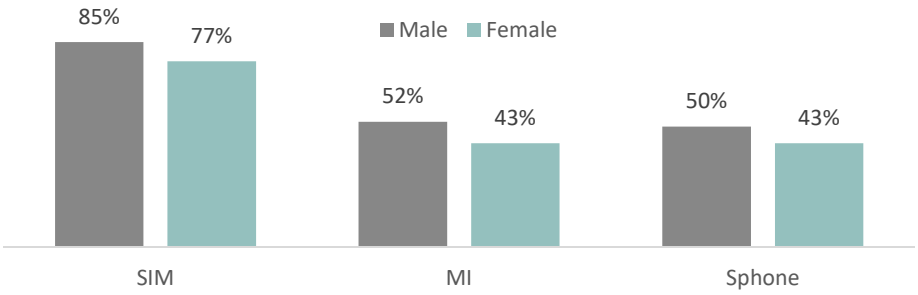


Chart 1 presents the percentage of the population that has adopted each of the three mobile technologies by gender. This includes all respondents in surveyed countries in each year, and each observation is given equal weight (i.e. there is no weighting by country population as in the GSMA Mobile Gender Gap Reports).

3 Empirical model

The empirical specification focuses on three dimensions of the use of mobile technology: (a) mobile ownership; (b) mobile internet use; (c) smartphone use. These dependent variables are binary and are defined as:

$$y_i \begin{cases} 1 \text{ if the } i\text{th individual has use of the mobile technology} \\ 0 \text{ if the } i\text{th individual does not have use of the mobile technology} \end{cases}$$

Table 1: description of dependent variables

Dependent variable	Description
Mobile owner	"Mobile phone owner" and "mobile owner" are used interchangeably in this paper to mean a person who has sole or main use of a SIM card, or a mobile phone that does not require a SIM, and uses it at least once a month.

[5] Aged 18+

[6] Gillwald et al. (2010) stratified their survey data into metropolitan, other urban and rural areas.

[7] The survey was delivered via interviewer-administered computer-assisted personal interviewing (CAPI). Survey interviews were conducted in the local language by both female and male interviewers. Interviews were conducted in respondents' homes, with the exception China where in-street interviewing was conducted.

Mobile internet	A "mobile internet user" is a person who has used the internet on a mobile phone at least once in the last three months. Mobile internet users do not have to personally own a mobile phone, and therefore can be non-mobile phone owners who use mobile internet by accessing it on someone else's mobile phone.
Smartphone	A smartphone user is someone with a smartphone that they have the sole or main use of. Respondents that don't have an active SIM are removed i.e. if they aren't a mobile owner following the description above.

We construct a model of the probability of adopting certain mobile technologies (mobile phone ownership (sim_i), mobile internet (mi_i), and smartphone use (spo_i)), conditional upon a vector of independent variables with gender (g_i) as the variable of interest, and a suite of control variables $(\boldsymbol{X})$.

We define a latent independent variable (y_i^*) as:

$$y_i^*(g_i, \boldsymbol{X};\ \beta_0, \beta_1, \boldsymbol{\beta}) \equiv \beta_0 + \beta_1 g_i + \boldsymbol{\beta X} + \varepsilon_i$$

This latent variable will determine the outcome variable (y_i) for each individual i in the following way:

$$y_i = \begin{cases} 0 \text{ if } y_i^* < 0 \\ 1 \text{ if } y_i^* \geq 0 \end{cases} \textit{where } y_i \begin{cases} sim_i \\ mi_i \\ spo_i \end{cases}$$

By making the assumption (amongst others) that the errors follow a logistic distribution (*i.i.d.*), we can use the logit model to estimate the conditional probabilities of the three mobile adoption outcomes:

$$Pr(y_i = 1 | g_i, \boldsymbol{X})$$

The independent variable of interest is a binary variable representing gender. As this variable is nominal, either 0 or 1 is assigned to male or female, and in this case female is represented by '1' and male by '0'. This should reflect the effects of non-observable aspects related to gender.

This is regressed alongside the following control variables of observable drivers of mobile adoption (Table 2).

Table 2: Variables used to specify the model

Socio-economic/demographic factor	Variable	Description
Gender	*Gender*	1 "Female", 0 "Male"
Age	*Age*	1 "18-24", 2 "25-34", 3 "35-44", 4 "45-54", 5 "55-64", 6 "65+"
Working status	*Not_working*	1 "Not working", 0 "Working"
Education	*Education*	1 "Primary or below", 2 "Secondary", 3 "Degree or above"
Rural-Urban Location	*Rural*	1 "Rural", 0 "Urban"
At least one child	*One_Dependent*	1 "At least one dependent", 0 "No dependents"
Literacy		1 "Low/no literacy skills", 0 "Fair/good literacy skills"

Household Income	*HH_Income*	1 "Low", 2 "Medium", 3 "High"[8]
Income Perception	*Income_Perception*	1 "Living comfortably on present income", 2 "Coping on present income", 3 "Finding it difficult on present income", 4 "Finding it very difficult on present income"
Years	*Year*	2018, 2019, 2017 (base)
Country	*Country*	Dummy variables for each country[9]

Given that different segments of female populations face different degrees of digital exclusion (see for example After Access, 2018), we also explore interactions between gender and the control variables in the model, for example with rural/urban location, income and education. This would show, for example, whether women living in rural areas face constraints to owning a phone or using mobile internet that are additional to those faced by women or rural populations more generally.

Network effects that influence mobile phone use can arise through membership of religious, cultural, economic and other communities (see for example Forenbacher et al., 2019). In our models, standard errors are clustered at a within country 'region' and 'geography' level (e.g. Lagos-Urban and Lagos-Rural) to take account of these network effects. Annex 6 presents the models at different levels of clustering.

4 Initial results

The results presented in Table 3 show that the average marginal effect[10] of 'female' on the probability of mobile ownership is -5 percentage points (pp), meaning that on average in our sample, women are 5pp less likely to own a phone then men (other factors being equal).[11] The marginal effects for mobile internet and smartphone adoption are -6pp and -4pp respectively. This suggests that even when other drivers of the gender gap are controlled for (particularly employment, income and education), women are still less likely to use mobile technology than men.[12]

The results also show that individuals in rural areas are 3pp less likely to own a phone than urban populations. The marginal effect is even greater for mobile internet (9pp) and smartphone adoption (8pp). Other control variables are in the expected direction and are statistically significant:

[8] For the 2019 and 2018 consumer survey, bands were: Low: Less than £15,000 a year, Medium £15,000 – £44,999 a year High: £45,000 or more. In the 2017 consumer survey: Low: Less than £20,000 a year, Medium £20,000 – £49,999 a year High: £50,000 or more - this causes issues with this variable as it cannot be pooled consistently across the three years. See Annex 4 on income variable alternatives. We also ran the regressions excluding the income variable and the results for other coefficients did not materially change.

[9] See Annex 1

[10] We choose to present marginal effects rather than odd-ratios, as many argue that odds ratios are often unintuitive and generally present little indication of the magnitude of the effect of a variable (Karaca-Mandic et al., 2012). If the probabilities are low and close to zero then the odds ratios are approximately equivalent to the risk ratio, however they are unlikely to be that low in our data. Secondly, the odds ratios of interaction terms cannot be interpreted in the same way as singular variables, as the coefficient on the interaction term is the natural logarithm of the ratio of two odds ratios. Given the complexity of odds ratios, even without the addition of the ratio of odds ratios for an interaction term, this output and approach may be of little use.

[11] Marginal effects are calculated using the delta method at the means of the rest of the explanatory variables, and therefore may not hold when they are set to other values. "The standard error of the average marginal effect of female across all subjects in the sample is not equal to the standard error of the marginal effect of female evaluated at the means of the explanatory variables although the results may be numerically close", (Karaca-Mandic et al., 2012).

[12] Although it is true that without a fully saturated model it is possible that the unobservable effect measured by gender could also be capturing other unobserved mechanisms which aren't directly, but indirectly, related to gender, we have included an extensive range of socioeconomic and demographic variables to control for the drivers of mobile adoption identified in the literature (Annex 2).

- Individuals with lower incomes are less likely to use mobile, mobile internet and smartphones. For example, compared to high-income individuals, those earning a low income are 15pp less likely to use mobile internet and those earning middle income are 5pp less likely.
- Those not working are 4pp less likely to use each type of mobile technology than those that are employed.
- Individuals that have only completed primary education (or less) are 32pp less likely to use mobile internet than those with a degree or above. Those that have completed secondary education are 15pp less likely to use mobile internet.
- Individuals with low levels of literacy are 7pp less likely to own a phone than those with good literacy skills
- Adoption of mobile technology generally declines with age. For example, individuals aged 25-34 are 6pp less likely to use mobile internet than those aged 18-24; those aged 35-44 are 15pp less likely; those aged 45-54 are 25pp less likely and those aged 65 or above are 43pp less likely[13]. For mobile ownership, however, the impact of age does not appear to start until 45 and above.
- Having dependents is not linked to higher or lower use of mobile internet or smartphones, but there is evidence that it is linked to higher mobile ownership.
- Annex 7 presents the results of regressions which also include a variable that represents if a person has a disability, and this variable interacted with gender. This survey question was limited to the 2019 survey only, and a smaller sample of countries. Although of high importance, this variable isn't included in the main regressions in order to take advantage of the full sample size by pooling the survey data where variables are available across the three years.

Table 3: Marginal effects on the conditional probability of mobile ownership[14]

	SIM	Mobile internet	Smartphone
Female	-0.0548***	-0.0620***	-0.0429***
	(-10.93)	(-13.62)	(-9.90)
Rural	-0.0347***	-0.0897***	-0.0821***
	(-5.21)	(-13.22)	(-12.76)
Not working	-0.0395***	-0.0369***	-0.0397***
	(-8.16)	(-8.45)	(-8.79)
Low income	-0.0562***	-0.149***	-0.164***
	(-6.34)	(-17.35)	(-18.71)
Middle income	-0.0154*	-0.0534***	-0.0668***
	(-2.14)	(-7.27)	(-8.79)
One dependent	0.00922*	0.00284	0.00309
	(2.18)	(0.65)	(0.65)
Low education	-0.135***	-0.319***	-0.303***
	(-19.85)	(-34.42)	(-31.44)
Middle education	-0.0556***	-0.145***	-0.141***
	(-9.29)	(-19.31)	(-18.67)
No/low literacy	-0.0738***	-0.134***	-0.123***
	(-12.29)	(-18.20)	(-18.26)

[13] This is likely to be a function of experience using relatively novel mobile technology in youth, and the younger generations with this experience will not necessarily give up this adoption when older.

[14] For a logit or probit model, the outcome of interest is the conditional probability that the dependent variable equals 1.

Age 25-34	0.0131**	-0.0590***	-0.0479***
	(2.91)	(-12.99)	(-9.06)
Age 35-44	0.00250	-0.151***	-0.114***
	(0.51)	(-26.23)	(-18.98)
Age 45-54	-0.0343***	-0.247***	-0.200***
	(-6.11)	(-38.27)	(-29.91)
Age 55-64	-0.0842***	-0.351***	-0.296***
	(-11.78)	(-46.32)	(-33.83)
Age 65+	-0.147***	-0.431***	-0.363***
	(-16.55)	(-50.02)	(-39.59)
N	61,526	61,526	59,457

* p<0.1, ** p<.05, *** p<.01. Z-scores in brackets.

4.1 Regional analysis

In Annex 5, we present the marginal effects when restricting the sample by region[15],[16], and in the mobile technology adoption rates by gender in each surveyed country are presented. In Table 4 below, the marginal effects of gender are presented for each region. We find that when controlling for other relevant factors:

- Women are as likely as men to use mobile technology (either mobile, mobile internet or smartphones) in Latin America and the Caribbean, once other relevant factors are controlled for.
- In the other regions, there is an unobservable gender effect – this is greatest in South Asia, where women are 20pp less likely to own a mobile phone than men, followed by Sub-Saharan Africa, and then MENA and East Asia & Pacific

Table 4: Marginal effects on the conditional probability of mobile technology adoption, by region

	East Asia & Pacific			Latin America & Caribbean			Middle East & North Africa		
	SIM	MI	Sphone	SIM	MI	Sphone	SIM	MI	Sphone
Female	-0.03***	-0.04***	-0.02*	-0.003	0.01	0.004	-0.03**	-0.05***	-0.05**
	South Asia			**Sub-Saharan Africa**					
	SIM	MI	Sphone	SIM	MI	Sphone			
Female	-0.20***	-0.15***	-0.14***	-0.04***	-0.08***	-0.03***			

* p<0.1, ** p<.05, *** p<.01

4.2 Interaction variables

Including interactions between variables can inform how the marginal effect of gender changes with other variables – for example when the urban/rural indicator changes. For each observation, the cross-derivative will be the change in the conditional probability that y=1 for a change in rural and gender. It is the difference between male and female of the marginal effect of rural on the conditional probability that y=1.

In Table 5 below, we have calculated the derivative of the gender variable at different values of the 'urban/rural' variable using the mobile ownership regression as an example. The effect of the interaction term is calculated by taking the difference of these marginal effects (-4 percentage points). This represents the effect of the statistical change between male and female in an urban environment and the effect of the change

[15] East Asia & Pacific, Latin America & Caribbean, Middle East & North Africa South Asia, and Sub-Saharan Africa

[16] In these initial regressions, countries are unweighted, however there is implicit weight given to India and China based on their populations, as at least 1,000 interviews a year were conducted in each country surveyed, with 2,000 conducted in India and China.

between male and female in a rural environment (Karaca-Mandic et al., 2012). In other words, it shows that women are 4pp less likely than men to own a mobile phone in rural areas compared to urban areas.

Table 5: The marginal effect of 'rural' and 'female' on the conditional probability of mobile ownership[17]

Marginal effect	dy/dx
Of Gender in Urban	-0.0336
Of Gender in Rural	-0.0719
Difference (Female#Rural)	*-0.0383*

Table 6 presents the derivatives of gender at different states of the other socioeconomic and demographic variables in the regression. As explained above, taking the differences of the marginal effects at each step will result in the marginal effect of the interaction term. In summary, the results show in addition to women being – on average – less likely to own a phone or mobile internet than men, this effect is 'enhanced' (or made worse) by the following factors:

- Living in rural areas.
- Unemployment (e.g. women are 2.5pp less likely to use mobile internet than men if they are unemployed, compared to if they are employed)
- Having lower levels of literacy (e.g. women are 3.4pp less likely than men to own a mobile phone if they have low levels of literacy, compared to if they have high literacy levels)
- Having dependents (this implies that the positive effect of having a dependent for mobile phone ownership is only for men).

Table 6: the marginal effect of gender at different states of the control variables included in the model

	Rural				Not working		
	SIM	**MI**	**Sphone**		**SIM**	**MI**	**Sphone**
Urban	-0.0336***	-0.0499***	-0.0290***	**Working**	-0.0494***	-0.0520***	-0.0351***
Rural	-0.0719***	-0.0790***	-0.0607***	**Not working**	-0.0621***	-0.0765***	-0.0544***
	One dependent				**No/low literacy**		
	SIM	**MI**	**Sphone**		**SIM**	**MI**	**Sphone**
No dependents	-0.0337***	-0.0503***	-0.0354***	**Literate**	-0.0463***	-0.0621***	-0.0389***
At least 1 dependent	-0.0641***	-0.0675***	-0.0465***	**No/low literacy**	-0.0806***	-0.0861***	-0.0787***

* p<0.1, ** p<.05, *** p<.01

4.3 Isolating the impact of unobservable and observable effects

The three charts below use the predicted probabilities from the three logit models (mobile ownership, mobile internet use and smartphone ownership) to isolate the effect of observable and unobservable drivers of the gender gap. Chart 1 shows that across the sample (for all countries and in all years), 85% of men and 77% of women owned a mobile phone respectively[18]. If we set the gender variable to zero, in order to remove any 'gender' effect, the difference between the predictions for 'female' and the predictions without the gender effect can be attributed to the non-observable aspects related to gender, such as discrimination and cultural norms. The effect of these unobservable aspects reduces the predicted probability of mobile ownership by

[17] For a logit or probit model, the outcome of interest is the conditional probability that the dependent variable equals 1.
[18] These predicted probabilities are equal to observed values in the dataset (Chart 1).

women by 6 percentage points. The remaining difference of 2 percentage points can be attributed to observable characteristics, such as income and education.

Similarly, we find that unobservable factors are significant drivers of the gender gap in mobile internet use and smartphone adoption, reducing the predicted probabilities by 5 percentage points and 7 percentage points respectively.

Chart 2: predicted probability of mobile ownership by gender: observable and unobservable differences

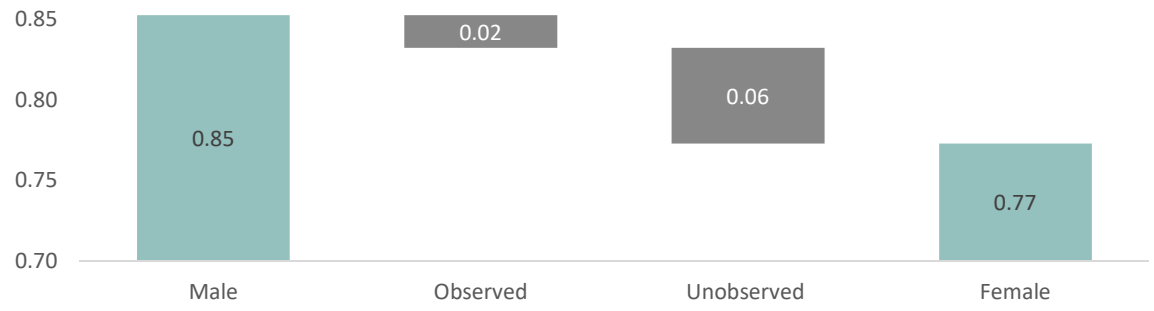
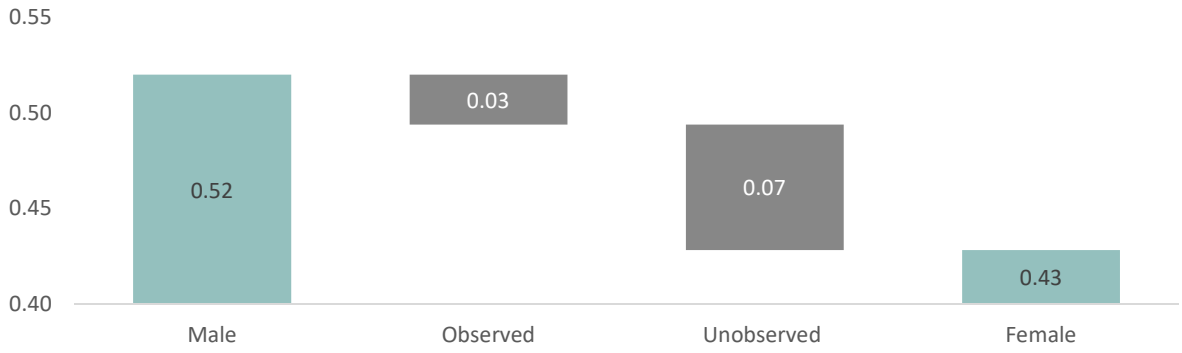


Chart 3: predicted probability of mobile internet adoption by gender: observable and unobservable differences

Chart 4: predicted probability of smartphone ownership by gender: observable and unobservable differences



5 Next steps

5.1 Additional analysis

We have initially explored the heterogeneity in the marginal effect of gender for regions, but we will also look at how the drivers of the gender gap vary by country. The GSMA Intelligence Global Consumer Survey includes questions on barriers to technology adoption that could be utilised for additional analysis on the mechanisms behind the gender gap[19].

Subsequent to the initial adoption of the different types of mobile technology, there may also be differences in the use of mobile technology by gender (see for example Girl Effect, Vodafone Foundation and MIT D-Lab, 2018) e.g. use of certain features or applications, such as mobile money accounts. As an extension to this work, we could replace the dependent variables to investigate the drivers of use of certain features of mobile technology.[20]

5.2 Identification and robustness improvements

We might be concerned about endogenous sample selection, as we can assume that an individual can only use mobile internet if they have access to a phone (SIM), and many of the same variables ($\boldsymbol{X}_i$) will influence mobile ownership (sim) as well as mobile internet use (mi). We observe not potential mobile internet use, but actual mobile internet use, and this is not representative of the whole population (because mobile phone access isn't random). The drivers of mobile internet adoption might not be the same once there is widespread access to mobile/SIM technology.

Allen et al. (2016) suggest using the Heckman selection model as they only observe when an agent makes a choice to use a bank account for savings when they have a bank account in the first place. The Heckman selection model could be used to estimate a relationship between gender and mobile internet adoption consistent with the whole population. The aim is to model both the relationship between gender and mobile internet adoption, and the mechanism by which individuals access mobiles/SIMs using a new variable (z_i)[21], with the assumption of independence between this new variable and mobile internet adoption (the exclusion restriction).

$$mi_i = mi(\boldsymbol{X}_i, sim_i, \varepsilon_i^*)$$

$$sim_i = sim(\boldsymbol{X}_{i,} z_i, \mu_1)$$

We might also be concerned about potential endogeneity between the binary dependent 'technology adoption' variables and the independent variables. For example, Aker and Mbiti (2010) set out mechanisms through which mobile phones can provide economic benefits to consumers and producers in Sub Saharan Africa. One method to address this is the use of instrumental variables. We could also consider explanatory variables that make use of the time variation in survey data, however this would need to exploit exogenous changes over time e.g. technological shocks or policy changes. For the moment, these approaches are outside of the scope of this study as our primary objective is to understand the factors that are strongly associated with mobile ownership, mobile internet use and smartphone ownership.

[19] For example, exploiting the Blinder-Oaxaca decompositions to identify how much each factor explains the gender gap, or using cross regional variation in gender gaps together with demand and financial factors.

[20] In the GSMA Intelligence Consumer Survey, a number of questions are asked of consumers on what activities they carry out on their mobile phones. This includes: using mobile banking/mobile money apps; purchasing goods/services; using healthcare services; educational services; e-Govt services; visit social networking websites; play games on a mobile phone; listen to music using a mobile phone.

[21] Potentially supply-side factors (e.g. coverage) could drive access to SIM only adoption.

In addition to a probability model, as a robustness check[22] we could consider applying the Ñopo (2008) matching method[23] used in the After Access (2018) paper. It allows us to quantify the explained and unexplained drivers of the gender gap. The wage gap decomposition in Ñopo's paper compares groups of observations with similar discrete and observable characteristics. The aim is to obtain a partitioned data set that contains four groups of observations of matched and unmatched males and females based on the characteristics (e.g. the control variables set out above). The observations within the sets of matched males and females have the same empirical distributions of probabilities for their characteristics, and the method is based on the assumption that individuals with the same observable characteristics should have the same level of technology adoption (in the case of Ñopo's paper, receive the same wages) regardless of their gender (Ñopo, 2008).

Following this methodology, the mobile technology gap (Δ) can be explained as:

$$\Delta = (\Delta_M + \Delta_X + \Delta_F) + \Delta_0$$

Whereby:

Δ_M: (all males) is the part of the technology gap that can be explained by differences between two groups of males, i.e. those who have characteristics that can be matched to female characteristics and those who do not.

Δ_X: (females and males) is the part of the technology gap that can be explained by differences in the distribution of characteristics between males and females.

Δ_F: (all females) is the part of the gap that can be explained by the differences in characteristics between two groups of females, i.e. those who have characteristics that can be matched to male characteristics and those who do not.

Δ_0: (females and males) the part of the technology gap that cannot be attributed to differences in characteristics. This difference can be attributed to unobservable characteristics.

To summarise, three of them $(\Delta_M + \Delta_X + \Delta_F)$ can be attributed to the existence of differences in individuals' characteristics, and the forth Δ_0 to differences in unobservable characteristics (Ñopo, 2008).

[22] The Ñopo methodology does not solve the selection bias without assuming ignorability, therefore, as another extension, non-parametric methodologies that try to account for selection bias when applying decompositions could be explored.
[23] Stata command: 'nopomatch'.

Annexes

Annex 1: List of countries and years in the GSMA Intelligence consumer survey

2017/18 (28 countries)	2018/19 (18 countries)	2019/20 (15 countries)
Algeria	Algeria	Algeria
Argentina	Argentina	Bangladesh
Bangladesh	Bangladesh	
Brazil	Brazil	Brazil
Cameroon		
Chile		
China	China	
Colombia		
Côte d'Ivoire	Côte d'Ivoire	
Dominican Republic	Dominican Republic	
Egypt		
Ghana		
Guatemala	Guatemala	Guatemala
India	India	India
Indonesia	Indonesia	Indonesia
Kenya	Kenya	Kenya
Mexico	Mexico	Mexico
Morocco		
	Mozambique	Mozambique
Myanmar	Myanmar	Myanmar
Nicaragua		
Nigeria	Nigeria	Nigeria
Pakistan	Pakistan	Pakistan
Philippines		
		Senegal
Sierra Leone		
South Africa	South Africa	South Africa
Tanzania	Tanzania	
Thailand		
		Uganda
Vietnam		

Annex 2: Literature review summary

Title	Background	Scope	Results
Determinants of mobile phone ownership in Nigeria, (Forenbacher et al., 2019).	The digital divide is the "gap between individuals, households, businesses and geographic areas at different socio-economic levels with regard to both their opportunities to access information and communication technologies (ICTs) and to their use of the Internet for a wide variety of activities" (OECD, 2018). To address the digital divide, the problem must be approached from multiple perspectives. Srinuan and Bohlin (2011) as well as Helbig, Gil-Garcia, and Ferro (2009) suggest identifying determinants of the digital divide on three levels: (1) level of access (e.g. infrastructure), (2) multi-dimensional level (e.g. education/income/age) and (3) multi-perspective level (e.g. gender). The digital divide can also be influenced by network effects arising through membership in religious, cultural, economic and other communities.	**Data:** Nigeria (national) from Statistics South Africa (2012), which came from the Research ICT Africa Household and Small Business Access and Usage Survey 2011-2012 (Research ICT Africa, 2012). **Model:** Logit model with binary dependent variable: mobile phone ownership. **Independent variables:** Age, income, rural/urban, electricity, gender, education, employment, literacy in mother tongue, literacy in English, informal work, social capital.	Their result was that the female gender was associated with lower likelihood of owning a mobile phone (coefficient, −0.474). The results also suggest that factors such as geographic location and income may not strongly influence mobile phone ownership The strongest factors appeared to be education level, informal work, social engagement, type of electricity supply and employment status. Therefore, policy makers should target younger adults, provide training in digital literacy specifically for mobile phone use, invest in electricity supply', and develop content and applications in non-English languages.
Understanding the Gender Gap in the Global South, (After Access, 2018).	The After Access 2018 report includes a comparative assessment of ICT access the Global South, and three regional studies. The extent of mobile phone ownership (and the gender gap) broadly aligns with gross national income (GNI) per capita. It is also noted in the report that most research fails to assess the intersectional nature of marginalisation (class, race). The regional study in Africa highlights that education and low income to be the main determinant(s) of access to the internet, and women are	The regional Latin America study is the most relevant for this paper: **Data:** Argentina, Colombia, Guatemala, Peru, Paraguay (After Access Survey, 2017). **Independent variables:** age, education level, number of children in the household, location., language, SEC and occupation **Model:** Adapting methodology of Ñopo	**The explained gender gap components:** Age (negative effect), education (positive effect), children (positive effect), rural & local language (negative effect), SEC (positive effect). **The unexplained gender gap components:** discrimination, cultural factors, sexism, racism etc.

	generally less educated, less employed and have lower income than men. The regional study in Latin America identifies the factors that determine the ICT gender gap.	(2008) (a non-parametric alternative to the Blinder-Oaxaca (BO) wage gap decomposition) to estimate the effect of observable (e.g. education) and non-observable factors (e.g. discrimination) on IT adoption.	Recommendations are, for example, to focus on improving educational and labour opportunities for women to reduce ICT gender gaps in Argentina, Colombia and Paraguay. However, for Peru and Guatemala, changing gender stereotypes and sexism is more urgent.
The foundations of financial inclusion: Understanding ownership and use of formal accounts, (Allen et al., 2016).	This paper examines the extent to which different individual characteristics and country level variables and policies are associated with financial inclusion, but also exploits the unique individual level characteristics.	**Data:** 2011 Global Findex database for 123 countries and over 124,000 individuals. **Model:** Probit models with binary dependent variables for bank account use. **Individual level independent variables:** gender, income, rural/urban, education, married, employment **Country level independent variables:** GDP per capita, costs of opening/running bank accounts Also examine reported subjective barriers to financial inclusion. we create a binary variable that takes the value 1 if a respondent without a bank account confirms it as a barrier to having an account and 0 otherwise.	They are exploring the drivers of financial inclusion which are similar to those for mobile/mobile internet. This study also looks at different phases of financial inclusion, similar to the three dependent variables (stages) for mobile adoption. Also somewhat relevant for our study, or any extensions to this study, is that they "acknowledge endogeneity biases could be present in the estimations discussed so far since the distribution of financial inclusion outcomes could influence policies and many factors could jointly affect policies and financial inclusion." They conduct instrumental variable estimations (and use LPM) of the likelihood of owning an account and use the adoption of policies in the neighbouring countries and an instrument for the policies in place in a country – the assumption is that countries are likely

			to replicate the policies of their neighbours.
Comparing the Determinants of Internet and Cell Phone Use in Africa: Evidence from Gabon, (Penard et al., 2012)	"Compare the determinants and hindrances of both Internet and cell phone use in Gabon, using household survey data to explain and compare the determinants of Internet and cell phone penetration". They considered the following binary variables: "Internet use during the previous 3 months" (Internet) and "ownership of at least one cell phone" (Mobile).	**Data:** Gabonese survey relative to individual use of ICT. Responses were recorded face-to-face in the cities of Libreville and Port-Gentil, from July 1st to November 30th, 2008. **Model:** Bivariate probit model: This model jointly estimates the decisions to adopt the two information technologies. **Independent variables:** socioeconomic characteristics (gender, age, level of education, marital status, occupation and lifestyle), his/her ICT skills and experience and his/her social environment (friend/family internet use, member ship of a formal association or a savings group).	"Primary factors stimulating Internet use consist of a high level of education and computer skills. As regards cell phone use, the main obstacles are economic. Finally, an individual's age has a positive impact on cell phone use and a negative impact on Internet use." No significant gender effect for mobile, but significant gender effect for internet.
Gender Assessment of ICT Access and Usage in Africa (Gillwald, Milek, & Stork, 2010)	"Analysis of the data also reveals other, perhaps overriding, inequities that point to the fact that poor women may have more in common with poor men in their own and other countries than with less marginalised women in their own and other countries". "The survey was stratified into metropolitan, other urban and rural areas. Enumerator areas (EAs) were sampled for each stratum using probability proportional to size (PPS) from the national census sample frame. for each country EA sample frames were constructed through listing all households within an EA. Households were	**Data:** The data that is being used for this study stems from surveys conducted by Research ICT Africa (RIA) in 17 African countries during the end of 2007 and the beginning of 2008. **Model:** Nationally representative probit model for each country using sampling weights (source: Chabossou et al. (2008)). **Independent variables:** income, age, gender, tertiary education, secondary education, primary education,	"The results suggest that the variables used to analyse mobile phone adoption contribute significantly to explaining the probability of adoption. The gender variable is mostly insignificant as expected (11 countries). Only in Senegal, Tanzania and Burkina Faso does being a woman decrease the probability of mobile telephony adoption. In South Africa and Mozambique it even increases the probability significantly".

	then sampled using simple random sampling. The individual to be selected for each household was randomly selected from all household members and visitors that stayed at the home on the night the household was visited and that were 16 years of age or older".	vocational education, rural social network.	(Same data and same results as in: Chabossou, Stork, Stork, & Zahonogo, Mobile Telephony Access and Usage in Africa, 2009).

Annex 3: Robustness checks

The first three columns in the table below presents the marginal effects of the independent variables estimated using the original logit regression, compared to a probit regression, and a linear probability model. Looking at the significance, and similarity in magnitude and direction of the marginal effects, our findings are robust to different model specifications.

Table: Alternative models to estimate the probability of the three types of mobile technology adoption

	Logit			Probit			Linear Probability Model		
	SIM	Mobile internet	Sphone	SIM	Mobile internet	Sphone	SIM	Mobile internet	Sphone
Female	-0.0548***	-0.0620***	-0.0429***	-0.0538***	-0.0610***	-0.0423***	-0.0577***	-0.0599***	-0.0409***
Rural	-0.0347***	-0.0897***	-0.0821***	-0.0351***	-0.0913***	-0.0831***	-0.0332***	-0.0942***	-0.0841***
Not working	-0.0395***	-0.0369***	-0.0397***	-0.0393***	-0.0380***	-0.0399***	-0.0381***	-0.0403***	-0.0432***
Low income	-0.0562***	-0.149***	-0.164***	-0.0568***	-0.148***	-0.164***	-0.0468***	-0.148***	-0.162***
Middle income	-0.0154*	-0.0534***	-0.0668***	-0.0164*	-0.0518***	-0.0660***	-0.00731	-0.0501***	-0.0611***
One dependent	0.00922*	0.00284	0.00309	0.00875*	0.00387	0.00349	0.00898*	0.00799	0.00796
Low education	-0.135***	-0.319***	-0.303***	-0.131***	-0.318***	-0.303***	-0.110***	-0.334***	-0.318***
Middle education	-0.0556***	-0.145***	-0.141***	-0.0529***	-0.143***	-0.140***	-0.0338***	-0.137***	-0.136***
No/ low literacy	-0.0738***	-0.134***	-0.123***	-0.0775***	-0.132***	-0.121***	-0.0999***	-0.113***	-0.103***
Age 25-34	0.0131**	-0.0590***	-0.0479***	0.0119**	-0.0595***	-0.0479***	0.00919*	-0.0654***	-0.0516***
Age 35-44	0.00250	-0.151***	-0.114***	0.00203	-0.151***	-0.114***	-0.00164	-0.157***	-0.118***
Age 45-54	-0.0343***	-0.247***	-0.200***	-0.0351***	-0.246***	-0.200***	-0.0361***	-0.249***	-0.200***
Age 55-64	-0.0842***	-0.351***	-0.296***	-0.0840***	-0.347***	-0.294***	-0.0869***	-0.345***	-0.289***
Age 65+	-0.147***	-0.431***	-0.363***	-0.151***	-0.424***	-0.359***	-0.176***	-0.402***	-0.345***

* p<0.1, ** p<.05, *** p<.01.

Annex 4: Variation on income variable: Income perception

The table below presents the results including an alternative variable for income: the income perception variable. We include this as an alternative measure, as the income variable buckets are not stable over the survey years, and the response to that survey question can be relatively low. However, the general direction of the coefficients are similar across the two income variables, and the other independent variables in the model remain stable – this further increases confidence in the robustness of the three models.

Table: Marginal effects of alternative variables for income (income brackets vs. income perception)

	Income		
	SIM	MI	Sphone
Low income	-0.0562***	-0.149***	-0.164***
	(-6.34)	(-17.35)	(-18.71)
Middle income	-0.0154*	-0.0534***	-0.0668***
	(-2.14)	(-7.27)	(-8.79)
	Income Perception		
	SIM	MI	Sphone
Finding it very difficult on present income	-0.0449***	-0.101***	-0.119***
	(-5.80)	(-11.63)	(-12.74)
Finding it difficult on present income	-0.0211**	-0.0633***	-0.0795***
	(-2.91)	(-8.84)	(-9.93)
Coping on present income	-0.00547	-0.0203***	-0.0328***
	(-0.81)	(-3.29)	(-4.56)
	With 'female' interaction term		
	Income		
	SIM	MI	Sphone
Low	-0.0527***	-0.0589***	-0.0399***
	(-10.04)	(-10.30)	(-7.53)
Middle	-0.0606***	-0.0695***	-0.0510***
	(-7.47)	(-10.34)	(-7.31)
High	-0.0499***	-0.0735***	-0.0471***
	(-4.42)	(-6.03)	(-3.66)
	Income Perception		
	SIM	MI	Sphone
Finding it very difficult on present income	-0.0497***	-0.0611***	-0.0306**
	(-5.68)	(-6.16)	(-2.95)
Finding it difficult on present income	-0.0500***	-0.0583***	-0.0335***
	(-6.49)	(-7.95)	(-4.41)
Coping on present income	-0.0452***	-0.0513***	-0.0338***
	(-6.98)	(-7.58)	(-4.82)
Living comfortably on present income	-0.0530***	-0.0656***	-0.0498***
	(-5.74)	(-5.49)	(-4.51)

* p<0.1, ** p<.05, *** p<.01. Z-scores in brackets.

Annex 5: Global regional differences

The marginal effects, when restricting the sample by region, for all the socioeconomic and demographic variables are presented below (as part of regressions that also include interaction terms).

	East Asia & Pacific			Latin America & Caribbean			Middle East & North Africa			South Asia			Sub-Saharan Africa		
	SIM	MI	Sphone	SIM	MI	Sphone	SIM	MI	Sphone	SIM	MI	Sphone	SIM	MI	Sphone
Female	-0.0300***	-0.0397***	-0.0219*	-0.00282	0.0110	0.00405	-0.0266**	-0.0516***	-0.0476**	-0.197***	-0.154***	-0.144***	-0.0410***	-0.0813***	-0.0329***
Rural	-0.0276	-0.0948***	-0.0713***	-0.0159	-0.0761***	-0.0510***	-0.0367	-0.0915**	-0.0869**	-0.0422*	-0.0637***	-0.0678***	-0.0462***	-0.106***	-0.113***
Not working	-0.0300***	-0.0380***	-0.0529***	-0.0435***	-0.0486***	-0.0431***	-0.0151	-0.0325	-0.0192	-0.0208	-0.00799	-0.0115	-0.0344***	-0.0319***	-0.0282***
Low income	-0.0544***	-0.125***	-0.138***	-0.0738***	-0.152***	-0.203***	-0.0164	-0.107*	-0.124*	-0.0449	-0.153***	-0.139***	-0.0564***	-0.177***	-0.177***
Middle income	-0.0191	-0.0542***	-0.0664***	-0.0187	-0.0492***	-0.0898***	0.00506	-0.0232	-0.0357	-0.0129	-0.0573***	-0.0415*	-0.00589	-0.0654***	-0.0709***
One dependent	-0.0107	-0.00425	-0.00103	0.00465	0.0246**	0.0160	0.00533	0.0151	0.0265	0.0241	-0.00914	-0.0216*	0.0135	-0.0116	-0.000504
Low education	-0.123***	-0.305***	-0.278***	-0.144***	-0.244***	-0.253***	-0.0467**	-0.297***	-0.277***	-0.146***	-0.293***	-0.275***	-0.120***	-0.448***	-0.421***
Middle education	-0.0626***	-0.154***	-0.121***	-0.0601***	-0.0892***	-0.114***	-0.0168	-0.119***	-0.112***	-0.0412*	-0.165***	-0.158***	-0.0482**	-0.206***	-0.203***
No/low literacy	-0.0762***	-0.159***	-0.121***	-0.133***	-0.182***	-0.185***	-0.0559**	-0.166***	-0.163***	-0.0515***	-0.0566***	-0.0636***	-0.0891***	-0.151***	-0.126***
Age 25-34	-0.0143	-0.0790***	-0.0531***	0.0128	-0.0324***	-0.0387***	-0.0167	-0.0517**	-0.0716***	0.0360**	-0.0711***	-0.0624***	0.0203**	-0.0629***	-0.0346***
Age 35-44	-0.0351***	-0.200***	-0.131***	0.00268	-0.113***	-0.101***	-0.0103	-0.141***	-0.139***	0.0122	-0.167***	-0.137***	0.0231**	-0.141***	-0.0905***
Age 45-54	-0.0848***	-0.330***	-0.242***	-0.0198*	-0.196***	-0.194***	-0.0429**	-0.255***	-0.233***	-0.0253	-0.241***	-0.221***	-0.0178	-0.233***	-0.147***
Age 55-64	-0.134***	-0.471***	-0.382***	-0.0585***	-0.350***	-0.307***	-0.0439*	-0.363***	-0.335***	-0.143***	-0.309***	-0.296***	-0.0396**	-0.270***	-0.197***
Age 65+	-0.208***	-0.558***	-0.462***	-0.134***	-0.499***	-0.438***	-0.0783**	-0.446***	-0.441***	-0.217***	-0.323***	-0.301***	-0.0863***	-0.337***	-0.249***

* p<0.1, ** p<.05, *** p<.01.

Technology adoption by gender, country, and year (GSMA Intelligence Consumer Survey)

	SIM			Mobile Internet			Smartphone		
Country and Gender	2017	2018	2019	2017	2018	2019	2017	2018	2019
Algeria									
Male	86%	92%	92%	61%	62%	68%	51%	63%	69%
Female	82%	87%	86%	53%	57%	55%	45%	61%	58%
Argentina									
Male	85%	88%		72%	80%		64%	75%	
Female	87%	89%		78%	81%		68%	74%	
Bangladesh									
Male	83%	86%	86%	36%	31%	35%	37%	35%	37%
Female	55%	59%	61%	13%	13%	16%	17%	16%	21%
Brazil									
Male	82%	84%	84%	67%	69%	72%	58%	65%	68%
Female	83%	87%	85%	64%	70%	74%	54%	66%	68%
Cameroon									
Male	95%			50%			45%		
Female	92%			39%			37%		
Chile									
Male	94%			75%			70%		
Female	95%			76%			71%		
China									
Male	97%	96%		77%	83%		82%	86%	
Female	97%	96%		75%	82%		83%	87%	
Colombia									
Male	91%			61%			56%		
Female	90%			59%			54%		
Dominican Republic									
Male	89%	85%		69%	73%		65%	69%	
Female	85%	85%		74%	80%		65%	77%	
Egypt									
Male	99%			67%			69%		
Female	97%			45%			46%		
Ghana									
Male	84%			40%			41%		
Female	81%			26%			33%		
Guatemala									
Male	82%	82%	79%	67%	70%	66%	65%	63%	60%
Female	77%	72%	71%	59%	56%	56%	57%	51%	51%
India									
Male	85%	81%	79%	26%	37%	41%	23%	35%	39%
Female	64%	61%	63%	8%	17%	21%	9%	16%	16%
Indonesia									
Male	73%	73%	81%	32%	43%	51%	28%	45%	53%

Female	65%	65%	73%	29%	38%	44%	27%	38%	48%
Ivory Coast									
Male	90%	90%		39%	51%		37%	47%	
Female	77%	81%		23%	26%		32%	33%	
Kenya									
Male	92%	86%	92%	49%	43%	49%	38%	38%	44%
Female	85%	80%	87%	31%	26%	33%	28%	27%	34%
Mexico									
Male	85%	86%	86%	68%	69%	72%	59%	64%	63%
Female	81%	86%	84%	62%	70%	73%	52%	62%	64%
Morocco									
Male	97%			37%			36%		
Female	94%			40%			40%		
Myanmar									
Male	77%	86%	77%	49%	55%	56%	65%	75%	69%
Female	64%	75%	69%	32%	36%	42%	51%	58%	59%
Nicaragua									
Male	70%			47%			48%		
Female	61%			40%			42%		
Nigeria									
Male	88%	88%	89%	38%	49%	53%	32%	42%	48%
Female	84%	84%	82%	26%	35%	37%	24%	35%	38%
Pakistan									
Male	79%	79%	81%	27%	39%	38%	27%	41%	37%
Female	44%	50%	50%	11%	11%	20%	12%	12%	20%
Philippines									
Male	78%			45%			47%		
Female	79%			49%			42%		
Sierra Leone									
Male	72%			31%			28%		
Female	70%			26%			23%		
South Africa									
Male	90%	87%	89%	59%	63%	67%	50%	61%	68%
Female	90%	83%	83%	57%	53%	59%	51%	55%	59%
Tanzania									
Male	78%	87%		26%	35%		18%	28%	
Female	70%	74%		14%	16%		11%	15%	
Thailand									
Male	98%			63%			70%		
Female	97%			60%			67%		
Vietnam									
Male	98%			64%			62%		
Female	95%			60%			62%		
Mozambique									
Male		60%	59%		24%	33%		21%	26%

Female	46%	47%	10%	17%	13%	19%
Senegal						
Male		77%		52%		50%
Female		72%		40%		40%
Uganda						
Male		81%		21%		17%
Female		71%		14%		15%

Annex 6: Clustering approach

Standard errors are clustered at a within country 'region' and 'geography' level e.g. Lagos-Urban and Lagos-Rural. The models are relatively stable when using different clustering approaches. When clustering at a regional or country level, only the variables "one dependent" and "Age 25-34" change in level of significance, and only marginally (see table below).

Table: alternative levels of clustering – changes to model variables (only shown if changed in level of significance)

	Region-Geog			Region			Country		
	SIM	**MI**	**Sphone**	**SIM**	**MI**	**Sphone**	**SIM**	**MI**	**Sphone**
One dependent	0.00922*	0.00284	0.00309	0.00922*	0.00284	0.00309	0.00922	0.00284	0.00309
Age 25-34	0.0131**	-0.0590***	-0.0479***	0.0131**	-0.0590***	-0.0479***	0.0131*	-0.0590***	-0.0479***

* p<0.1, ** p<.05, *** p<.01.

Annex 7: Assistive technology: people with disabilities

Eight countries in the 2019 survey included a variable that indicated if a respondent was a person with a disability. Although of high importance, this variable isn't included in the main regressions in order to take advantage of the full sample size by pooling the survey data where variables are available across the three years.

The marginal effect of being a person with a disability is significant and negatively affects the probability of mobile ownership by -7 percentage points, and smartphone ownership by -9 percentage points (with no statistically significant link for mobile internet use). Looking at the intersection of gender and disability, the cross derivative shows that the effect of female and the effect of disability simultaneously reduces the probability of technology adoption by between -1 and -4 percentage points, with the most reduction in probability of technology adoption associated with those with a disability who are also female.

Table: Partial and cross-partial marginal effects of disability and disabilities with 'female' gender

	SIM	**MI**	**Sphone**
Person with disability	-0.0708**	-0.0147	-0.0945***
	With 'female' interaction term		
	SIM	**MI**	**Sphone**
Person with disability=0	-0.0631***	-0.0775***	-0.0620***
Person with disability=1	-0.0807**	-0.0850**	-0.0990*
Difference	*-0.02*	*-0.01*	*-0.04*

* p<0.1, ** p<.05, *** p<.01.

Control variables specified in Section 3 were included in the model

Annex 8: Identification and diagnostics

The table below presents includes the results of the logit regressions, both with and without interaction terms of the independent variables with gender[24]. Inspecting the p-values for the z-tests, we can see that the gender variable (female) is consistently significant across the models, either individually or as part of an interaction variable, and therefore improves the goodness-of-fit of the models for prediction of technology adoption.

Other diagnostics indicate that the regressions that include gender interaction terms are superior to those without. For example, the likelihood ratio tests that compare the model with and without interaction terms consistently produce p-values of less than 0.05, therefore we can conclude that the model with interaction terms fits the data significantly better than the model with only single variables. Similarly, the Wald tests for these models indicate that including these variables results in a statistically significant improvement in the fit of the model. In addition to this, the pseudo-R2 values improve with the addition of the interaction terms.

Table: logit model for prediction of mobile ownership (SIM), adoption of mobile internet and smartphone ownership – with and without interaction terms with gender (female)

	SIM		Mobile Internet		Smartphone	
Female	-0.424***	0.04	-0.413***	-0.433***	-0.267***	-0.191
	(0.044)	(0.173)	(0.035)	(0.123)	(0.030)	(0.125)
Rural	-0.274***	-0.132*	-0.586***	-0.495***	-0.504***	-0.409***
	(0.053)	(0.070)	(0.044)	(0.052)	(0.039)	(0.046)
Rural # Female		-0.251***		-0.192***		-0.202***
		(0.076)		(0.060)		(0.055)
Education: primary	-1.202***	-1.099***	-1.971***	-1.957***	-1.766***	-1.787***
	(0.080)	(0.105)	(0.058)	(0.070)	(0.056)	(0.067)
Education: Primary or below # Female		-0.199		-0.035		0.048
		(0.137)		(0.093)		(0.090)
Education: secondary	-0.589***	-0.474***	-0.925***	-0.969***	-0.838***	-0.857***
	(0.075)	(0.101)	(0.050)	(0.066)	(0.046)	(0.059)
Education: Secondary # Female		-0.226*		0.09		0.042
		(0.136)		(0.091)		(0.088)
Not working	-0.325***	-0.286***	-0.262***	-0.167***	-0.260***	-0.188***
	(0.038)	(0.050)	(0.029)	(0.041)	(0.028)	(0.040)

[24] All models also include year and country dummies in order to account for any variation, however the results do not include coefficients for these.

Not working # Female		-0.04		-0.169***		-0.128**
		(0.058)		(0.054)		(0.053)
Age: 25-34	0.103***	0.094*	-0.388***	-0.415***	-0.297***	-0.320***
	(0.038)	(0.057)	(0.029)	(0.042)	(0.032)	(0.044)
25-34 # Female		0.033		0.074		0.064
		(0.079)		(0.061)		(0.061)
Age: 35-44	0.019	-0.026	-0.957***	-1.066***	-0.690***	-0.734***
	(0.041)	(0.065)	(0.036)	(0.049)	(0.036)	(0.052)
35-44 # Female		0.082		0.236***		0.102
		(0.088)		(0.068)		(0.070)
Age: 45-54	-0.264***	-0.302***	-1.570***	-1.760***	-1.207***	-1.306***
	(0.044)	(0.064)	(0.043)	(0.059)	(0.042)	(0.059)
45-54 # Female		0.055		0.413***		0.219***
		(0.085)		(0.079)		(0.075)
Age: 55-64	-0.605***	-0.727***	-2.307***	-2.561***	-1.836***	-1.993***
	(0.050)	(0.072)	(0.057)	(0.075)	(0.060)	(0.075)
55-64 # Female		0.210**		0.570***		0.353***
		(0.091)		(0.099)		(0.094)
Age: 65+	-0.984***	-1.113***	-2.961***	-3.149***	-2.319***	-2.383***
	(0.055)	(0.067)	(0.074)	(0.091)	(0.068)	(0.082)
65+ # Female		0.226**		0.423***		0.136
		(0.096)		(0.110)		(0.112)
Income: low	-0.458***	-0.501***	-0.982***	-1.020***	-1.003***	-1.012***
	(0.078)	(0.094)	(0.056)	(0.073)	(0.054)	(0.071)
Income: low # Female		0.071		0.088		0.024
		(0.102)		(0.085)		(0.083)
Income: medium	-0.134**	-0.102	-0.358***	-0.369***	-0.412***	-0.398***
	(0.066)	(0.088)	(0.049)	(0.063)	(0.047)	(0.063)
Medium Income # Female		-0.062		0.031		-0.026
		(0.103)		(0.083)		(0.082)

At least one child	0.081**	0.214***	0.009	0.075**	0.016	0.053
	(0.033)	(0.044)	(0.029)	(0.038)	(0.030)	(0.037)
At least one child # Female		-0.254***		-0.115**		-0.07
		(0.052)		(0.051)		(0.050)
Low/no literacy skills	-0.538***	-0.460***	-0.868***	-0.785***	-0.755***	-0.629***
	(0.041)	(0.058)	(0.048)	(0.058)	(0.041)	(0.053)
Low/no literacy skills # Female		-0.139**		-0.192***		-0.273***
		(0.063)		(0.068)		(0.067)
Constant	4.164***	3.916***	4.012***	4.019***	3.366***	3.326***
	(0.201)	(0.217)	(0.115)	(0.133)	(0.124)	(0.139)
Observations	***61526***	***61526***	***61526***	***61526***	***59457***	***59457***
Pseudo-R2	***0.165***	***0.166***	***0.344***	***0.346***	***0.304***	***0.305***
Log-likelihood	***-24700***	***-24700***	***-27900***	***-27800***	***-28600***	***-28600***
Wald test p-value	***0.00***	***0.00***	***0.00***	***0.00***	***0.00***	***0.00***

* p<0.1, ** p<.05, *** p<.01. Standard errors in brackets.

Table: a selection of diagnostic tests for logit regressions

Diagnostic test	Purpose
Likelihood ratio test	Goodness-of-fit
Wald chi2 test	Goodness-of-fit
McFadden's pseudo-R2	Goodness-of-fit
Link test/omitted variable test	Specification – additional predictors (shouldn't be able to find other additional predictors except by chance if the model is correctly specified)
Box-Tidwell[25]	Transforms a predictor using power transformations and finds the best power for model fit based on maximal likelihood estimate.

[25] Initial results show no transformations, but this is to be revisited.

Likelihood ratio test

Likelihood ratio test comparing the model with and without interaction terms. The p-values for all three mobile technology categories are < 0.05, therefore we can conclude that the model with interaction terms fits the data significantly better than the model with only single variables.

	SIM	MI	Sphone
Likelihood ratio test: p-value	0.00	0.00	0.00

Link test

The predicted value from the model squared is only significant for the smartphone model (p-value < 0.05), this indicates that there may be omitted relevant variables – this will be investigated in further iterations of this paper.

	SIM		MI		Sphone	
	Coef.	P>z	Coef.	P>z	Coef.	P>z
_hat	0.97874	0.000	0.99886	0.000	0.997136	0.000
_hatsq	0.008169	0.343	-0.00696	0.098	-0.01671	0.000
_cons	0.006021	0.749	0.011829	0.352	0.025305	0.044

References

After Access (2018) Understanding The Gender Gap in The Global South. DIRSI, LIRNEasia, Research ICT Africa.

Ai, C., Norton, E., (2003) Interaction terms in logit and probit models, *Economics Letters*. 80 (1), pp. 123-129.

Aker, J. & Mbiti, I. (2010) Mobile Phones and Economic Development in Africa. *Journal of Economic Perspectives*. 24 (3), pp. 207-32.

Allen, F., Demirguc-Kunt, A., Klapper, L., Peria M S M., (2016) The foundations of financial inclusion: Understanding ownership and use of formal accounts, *Journal of Financial Intermediation*. 27, pp. 1-30.

Bimber, B., (2000) Measuring the Gender Gap on the Internet. *Social Science Quarterly,* 81 (3), pp. 868-876.

Buis, M., (2010) Stata tip 87: Interpretation of interactions in nonlinear models, *The Stata Journal.* [https://journals.sagepub.com/doi/pdf/10.1177/1536867X1001000211].

Cameron, A.C. & Trivedi, P.K., (2005) Microeconometrics: Methods and Applications. Cambridge University Press.

Chabossou, A., Stork, C., Stork, M. & Zahonogo. Z. (2009) Mobile Telephony Access and Usage in Africa. ICTD Conference Paper.

Forenbacher, I., Husnjak, S., Cvitić, I. & Jovović, I. (2019) Determinants of mobile phone ownership in Nigeria. *Telecommunications Policy*. 43 (7).

Girl Effect, Vodafone Foundation and MIT D-Lab, (2018) [https://static1.squarespace.com/static/5b8d51837c9327d89d936a30/t/5beaa1700e2e72ed39d21c5b/1542103477055/GE_VO_Full_Report_Digital.pdf]

Gillwald, A., Milek, A., & Stork, C., (2010) Gender Assessment of ICT Access and Usage in Africa, Research ICT Africa. 1 (5).

Greene, C., (2010) Testing hypotheses about interaction terms in nonlinear models, *Economics Letters.* 107 (2) pp. 291-296.

GSMA (2019), Connected Women The Mobile Gender Gap Report [https://www.gsma.com/mobilefordevelopment/wp-content/uploads/2019/03/GSMA-Connected-Women-The-Mobile-Gender-Gap-Report-2019.pdf]

GSMA (2019), State of Mobile Internet Connectivity Report [https://www.gsma.com/mobilefordevelopment/wp-content/uploads/2019/07/GSMA-State-of-Mobile-Internet-Connectivity-Report-2019.pdf]

Institute of Development Studies, (2018) Leaving No One Behind in a Digital World,

Karaca-Mandic, P., Norton, E., Dowd, B., (2012), Interaction Terms in Nonlinear Models, *Health Services Research.* 47 (1).

Kohler, U., Karlson, K., Holm, A., (2011), Comparing coefficients of nested nonlinear probability models, *The Stata Journal.* [https://www.stata-journal.com/sjpdf.html?articlenum=st0236]

McFadden, D. (1974) Conditional logit analysis of qualitative choice behavior. pp. 105-142 in P. Zarembka (ed.)

Norton, E., Wang, H., Ai, C., (2004), Computing interaction effects and standard errors in logit and probit models, *The Stata Journal*. [https://journals.sagepub.com/doi/pdf/10.1177/1536867X0400400206]

Penard, T., Poussing, P., Zomo Yebe, G. & Nsi Ella, P. (2012) Comparing the Determinants of Internet and Cell Phone Use in Africa: Evidence from Gabon. *Communications & Strategies.* 86, pp. 65-83.

Rodrik, D., (2008), The New Development Economics: We Shall Experiment, but How Shall We Learn, Harvard Kennedy School.

Schumacher, P., & Morahan-Martin, J. (2001). Gender, Internet and computer attitudes and experiences. *Computers in Human Behavior*, 17 (1), pp. 95–110.

Wooldridge, J., (2002), Econometric Analysis of Cross Section and Panel Data. The MIT Press.