Illiteracy and exclusion: Evidence from India

Nishant Chadha, Satyaki Chakravarty, and Soma Wadhwa

India Development Foundation

Abstract

This paper highlights the relationship between an old problem plaguing India, illiteracy, and newer developmental challenges, financial and digital exclusion. We use the Financial Inclusions Insight (FII) data from India on mobile phone ownership and access from 2013 - 2016 to establish that the use of mobile phones, banks and digital money increases with literacy levels. Illiterate women are significantly less likely to own and use a phone or a bank account compared to informally literate women who themselves are less likely to own and use a phone and bank account compared to formally literate women. These results are robust to inclusion of various controls and fixed effects. To allay concerns of omitted variables driving these results we supplement these results with those from primary data from a female adult literacy programme (FALP) from Uttar Pradesh in northern India. Using a mixed methods research design, we present both quantitative and qualitative evidence, that going through the programme increases both the quantity and quality of mobile phone use and use of bank account. However, this is still lower compared to formally literate women. These results assume significance in light of the fact that, in spite of a long history of adult literacy programmes, 50 per cent of women above the age of 15 in India continue to be illiterate (Census 2011). In fact, in the state of UP, where the programme is located, 56 per cent adult women are illiterate. The results of this paper show that purely supply side interventions will not be enough to bridge digital divides in terms of gender and education.

1 Introduction

"Digital technologies have spread rapidly in much of the world. Digital dividendsthe broader development benefits from using these technologieshave lagged behind. In many instances digital technologies have boosted growth, expanded opportunities, and improved service delivery. Yet their aggregate impact has fallen short and is unevenly distributed. For digital technologies to benefit everyone everywhere requires closing the remaining digital divide, especially in internet access. But greater digital adoption will not be enough. To get the

most out of the digital revolution, countries also need to work on the "analog complements"–by strengthening regulations that ensure competition among businesses, by adapting workers skills to the demands of the new economy, and by ensuring that institutions are accountable."(World Bank report 2016)

Digitisation can lower transaction costs, raise efficiency and improve labour productivity. It also promotes social and economic inclusion and its benefits go beyond pure economic returns. Governments across the world, but especially in developing countries have adopted e-Governance as a mantra for improving service delivery in often corrupt and inefficient public systems. India is no exception. Arun Jaitley, the Indian finance minister has claimed that the "JAM trinity as the potential to link all Indians into one common financial, economic and digital space"[1]. In spite of all these potential benefits however inclusion remains a big concern, with access to technology far from universal. According to the World Bank Report 2016, women 25 per cent less likely than men to have internet access.

This paper establishes the relationship between another aspect of deprivation, illiteracy, and digital and financial exclusion. Using nationally representative Financial Inclusion Insights (FII) data from 2013-2016, we show that illiterate women are less likely to own and use mobile phones and bank accounts compared to informally literate women who in turn are less likely to be 'included' compared to formally literate women. The results are robust to controlling for various individual and household characteristics. However, the concern still remains that unobserved variables may be driving these results.

To supplement the above findings and establish causality we use primary data from a female adult literacy programme (FALP). The programme ran in three districts of eastern Uttar Pradesh (UP) from 2013-2016. Specifically in six blocks of Mirzapur, Jaunpur and Sant Ravidas Nagar districts of eastern UP with 1,48,221 women participating in it. Our data collection was done in November, 2016. We find that the participation in the programme increases mobile phone ownership and usage significantly. It also increases the likelihood of having a bank account. However, despite this increase the use and access is lower than that of literate women in the same villages.

These results show that the challenges to digital inclusion do not always come from the 'digital world'.

2 Data

We use data from two separate sources, Financial Inclusion Insights (FII) and an adult literacy programme called Tara Akshar + (TA+) to test our hypotheses. A short description of each dataset are as follows:

Financial Inclusion Insights (FII): Intermedia and Bill and Melinda Gates Foundation (BMGF) partnered to start the FII programme. FII focuses on countries in Africa and Asia with a combined population of over 2 billion starting in the year 2013. InterMedia conducts annual nationally representative

[1] JAM refers to Jandhan, Aadhar and Mobile

surveys in various countries and is expanding. In this paper we use data for the following countries' data: Bangladesh, India, Indonesia[2], Kenya, Nigeria, Pakistan, Tanzania and Uganda for robustness checks and descriptive statistics, and only India for the main results and models which we present. Each year, InterMedia updates questionnaires and modifies definitions. Due to updation some variables become incomparable over years. Hence we use only those variables which are consistent over time, or can be modified to a uniform definition without convolving originality.

The sampling strategy followed by Intermedia in India is as follows: Sample from India is divided across towns and villages in the ratio of 31: 69 urban to rural respectively - similar to the universe population distribution of 33:67 urban to rural respectively. Strata based on population is created, called town classes and village classes and from each category samples are taken. Oversampling for few states is done to ensure adequate representation. Assam is treated as a separate state, other North-eastern Indian states are clubbed into one cluster. Jammu and Kashmir is not covered in the sample. The questionnaire is built on a modular approach with sections on mobile phones, banks, mobile money and non-bank financial institutions. Each module explores awareness, access and use. Post collection, the data is weighted to conform to the national population, and several algorithms are run to again confirm accuracy.

Each town and village is clubbed into town classes and village classes and are defined as follows:

- Town class 1: more than or equal to 40 lakh population
- Town class 2: 10 lakh to 40 lakh population, but not including 40 lakh
- Town class 3: 1 lakh to 10 lakh population, but not including 10 lakh
- Town class 4: 50 thousand to 1 lakh population, but not including 1 lakh
- Town class 5: less than 50 thousand population
- Village class 1: more than or equal to 3000 population
- Village class 2: 1000 to 3000 population, but not including 3000
- Village class 3: less than 1000 population

Tara Akshar + (TA+):

In this section we discuss the sampling design for data collection for the quantitative part of the study. In particular, we discuss:

- Sampling of treatment units for quantitative survey
- Sampling of control units for quantitative survey

[2]survey started in 2014 for Indonesia which is included in this study

Sampling of treatment units for quantitative survey: The quantitative survey was carried out for 1500 TA+ participants in 36 of the 435 programme villages. We followed a two stage stratified cluster sampling method to identify the respondents. We first stratified villages in each of the six intervention blocks into tertiles of the 2011 female literacy rate distribution. Then from each strata we selected two clusters (villages) for a total of 36 villages (2 clusters $\times$ 3 strata $\times$ 6 blocks).

Next, we chose about 40 TA+ participants from each cluster by another stage of stratified sampling. At the time of data collection TA+ classes were no longer running in any of the blocks. The GCs, however, were still active. Hence we stratified the population of TA+ participants in each cluster into five groups depending on how long ago they had finished the TA+ classes. The groups were as follows:

- Group 1 consists of participants who joined GC less than two months ago
- Group 2 joined the GC 5-6 months ago
- Group 3 graduated GC less than or equal to six months ago
- Group 4 graduated GC between six months to a year ago
- Group 5 graduated GC more than a year ago

Sampling of control units for quantitative survey: We defined the village to be the unit of intervention because, as mentioned earlier, the programme ran for multiple cycles in each village and the mobilisation for the programme was done at the community level. Thus a control group cannot be chosen from the programme villages because of a very strong possibility of contamination. Women who did not join the programme may still have been exposed to its messages because of the community mobilisation undertaken. We thus believe that the correct way to choose control units is to sample villages from neighbouring blocks where there was no presence of TA+.

To do so we drew up a list of villages in blocks that border the programme blocks and then from amongst these chose control villages through the propensity score matching (PSM). We chose 36 villages (each treatment village matched to a control village) from these neighbouring blocks that are the closest matches to our treatment villages. Note that this stage of PSM was carried out to select control villages from which to sample the control group of women. After identifying control villages we randomly sampled about 35 women from each village to form the control sample of women.

Villages were matched on the following variables obtained from the 2011 Census of India: population density, gender ratio, female literacy rate, proportion of scheduled caste, proportion of scheduled tribe, distance from sub-district head quarter (in km), distance from district head quarter (in km), nearest statutory town (in km), major district road status, other district road status, black topped (pucca) road status, gravel (kuccha) road status, all weather road status. This choice of variables is based on interviews with programme staff on

how the selection of treatment villages was done. Demographic characteristics and access to the village were most important.

To sum up, the sample for FII is composed of data on women in India for all the states barring Jammu and Kashmir for the years 2013 to 2016, i.e. four years and the sample for TA+ is composed of data on women in selected villages of western Uttar Pradesh for the year 2016.

2.1 Descriptive statistics

Table 2.1 and 2.1 reports descriptive statistics for FII and TA+ dataset respectively for all the variables we use in this paper. Both tables show difference in mean values of owning and access to mobile phones if literacy levels are different. Pairwise test of significant difference in means for the three literacy categories are shown in the appendix.

3 Econometric approach

FII dataset shows the relationship between literacy, use of mobile phones and use of financial services for women in India. Evidence from TA+ acts supplementary to the stated relationship, indicating causation, that is to say illiteracy causes digital and financial exclusion.

The variables of our interest for FII are:

- mobile ownership
- mobile access
- aware of mobile money
- use mobile money
- bank account ownership, and
- bank account access

which are taken as dependent variable in the various models reported.

The variables of our interest for TA+ are:

- mobile ownership
- mobile access, and
- bank account ownership

Each of the tables reported for FII contain five models pertaining to the stated dependent variables. Each model is represented by a column in a table. Each column is a model of the following specification:

$$y_{st} = Constant(\alpha) + \beta_1 Informally\_literate_{st} + \beta_2 Literate_{st} + \boldsymbol{X}_{\boldsymbol{st}}^{'}\gamma + \epsilon \quad (1)$$

where:

s = state

t = time

y_{st} = 1 if the woman says yes to the dependent variable; 0 otherwise

$Informally\_literate_{st}$ = 1 if woman is informally literate; 0 otherwise

$Literate_{st}$ = 1 if woman is literate; 0 otherwise

$\boldsymbol{X'_{st}}$ assumes all fixed effects and control variables:

$class$ = town class or village class fixed effects; categorised according to population for the town or village respectively

$state$ = state fixed effects

$state\_class$ = interaction between $state$ and $class$

$year$ = year fixed effects

$job\_current$ = 1 if the woman was engaged in a remunerative activity; 0 otherwise

edu_male = educational qualification of the male of the household

$household\_size$ = number of members in the household to which the woman belongs

$schoolgoing\_age\_child$ = number of school-going children in the house, if any

$farmland\_own$ = 1 if the woman's family owns farmland; 0 otherwise.

The columns are described as follows:

1. Column (1): dependent variable regressed on dummies for $Informally\_literate$ and $Literate$ keeping illiterate as base i.e. equation 1 without $\boldsymbol{X'_{st}}$
2. Column (2): column (1) with $year$, $class$ and $state$ fixed effects
3. Column (3): column (1) with $year$ and $state\_class$ interaction
4. Column (4): column (3) with controls common to all the years ($job\_current$ and edu_head)
5. Column (5): column (4) with all controls ($household\_size$, $schoolgoing\_age\_child$ and $farmland\_own$).

Robust standard errors are used for column (1) while standard errors clustered at $state \times class$ are used for columns (2) to (5).

Each of the tables reported for TA+ contain three models pertaining to the stated dependent variables. Each model is represented by a column in a table. Each column is a model of the following specification:

$$y = Constant(\alpha) + \beta_1 treatment + \boldsymbol{X}'\gamma + \epsilon \tag{2}$$

where:

$y = 1$ if the woman says yes to the dependent variable; 0 otherwise

$treatment = 1$ if woman receives treatment; 0 otherwise

$\boldsymbol{X}'$ assumes all control variables:

age = age of the woman;

$child\_num$ = number of children she had given birth to and brought them up;

$adult\_num$ = number of adult members in the household;

$income\_main$ = main source of income;

$members\_outside$ = members not residing in the household where the woman lives;

$job\_current = 1$ if the woman is engaged in a remunerative activity; 0 otherwise

The columns are described as follows:

1. Column (1): dependent variable regressed on *treatment* keeping only illiterate from the control sample as base i.e. equation 2 without controls
2. Column (2): column (1) with controls;
3. Column (3): column (2) using PSM (women with similar characteristics from treatment and control samples are matched);
4. Column (4): dependent variable regressed on *treatment* keeping only literate from the control sample as base i.e. equation 2 with controls

Robust standard errors are used for all models for TA+ regressions except the PSM model.

4 Results and Discussion

Tables 3, 4, 5, 6, 7 and 8 reports coefficients for informally literates and literates and compares them to the base i.e. illiterates using FII dataset. Columns (1) to (5) report OLS regression models.

As was indicated in the descriptive statistics and pairwise significance tests, the results hold true, after controlling for various factors mentioned and are fairly consistent over all the models. As literacy increases, probability of owning and accessing mobile phones increases, probability of awareness of mobile money also increases with increase in literacy. However, use of mobile money, owning bank account and access to bank account does not significantly increase with informal literacy and only increases with formal literacy. Indicating a limitation of informal literacy, it can definitely improve inclusion, but is limited in certain aspects, which requires formal literacy.

The exact trend is seen in the TA+ data as well, reported in Tables 9, 10 and 11. Here, columns (1), (2) and (4) report OLS regression models. Column (3) uses PSM technique to match women of similar characteristics (characteristics are defined as the control variables) and reports results for those. Illiterate women in control sample are compared to the neo-literates, who have just received the knowledge through the literacy programme, which shows neo-literate women use and have access to mobile phones significantly more than the illiterate women. While benefits from informal literacy can loosely be compared to benefits from TA+, if we assume TA+ is similar to other adult literacy programmes, it helps us in explaining that the trend we see for entire India is exactly the trend we see in TA+.

A Robustness checks

A.1 pairwise t-test

Table 1: Descriptive statistics by literacy level (FII dataset)

Variable	**Illiterate**		**Informally literate**		**Literate**	
	mean	**sd**	**mean**	**sd**	**mean**	**sd**
Own mobile	0.25	0.43	0.38	0.49	0.62	0.48
Access to mobile	0.59	0.49	0.61	0.49	0.66	0.47
Bank A/C	0.48	0.50	0.51	0.50	0.63	0.48
Access to Bank A/C	0.01	0.12	0.02	0.13	0.03	0.17
Mobile money aware	0.00	0.07	0.01	0.10	0.10	0.30
Use mobile money	0.00	0.04	0.00	0.04	0.02	0.13
Household size	4.41	2.35	4.42	2.16	4.52	2.20
School-going children	1.90	1.40	1.86	1.30	1.68	1.22
Male general Education	2.22	0.75	2.41	0.77	3.20	1.01
Currently employed	0.46	0.50	0.53	0.50	0.48	0.50
Basic phone	0.89	0.32	0.85	0.36	0.72	0.45
Feature phone	0.62	0.49	0.54	0.50	0.49	0.50
Smart phone	0.52	0.50	0.40	0.49	0.35	0.48
Own sim card	0.18	0.38	0.29	0.45	0.48	0.50
Own farmland	0.44	0.50	0.42	0.49	0.40	0.49

Note: Variables own mobile, access to mobile, bank A/C, access to bank A/C, mobile money aware, use mobile money, currently employed, basic phone, smart phone, own sim card and own farmland take value 1 if the answer is yes to the question and take value 0 otherwise.

Variables household size and school-going children take absolute values.

Variable male general education reports the general education of male member of the household; husband is preferred to any other male member of the household.

The variable takes value 1 if no male member is present;
takes value 2 if illiterate;
takes value 3 if education till middle school;
takes value 4 if education till secondary school; and
takes value 5 if diploma, graduate or post-graduate and above.

Table 2: Descriptive statistics by literacy level (Tara Akshar +)

Variable	**Illiterate**		**Neo-literate**		**Literate**	
	mean	**sd**	**mean**	**sd**	**mean**	**sd**
Own Mobile	0.43	0.50	0.55	0.50	0.71	0.46
Access to mobile	0.83	0.38	0.91	0.28	0.92	0.27
Number of adults	3.76	2.32	4.45	3.11	4.50	3.37
Number of children	3.19	1.91	3.34	3.35	3.04	2.29
Income source	3.13	1.38	3.03	1.54	2.88	1.44
Age	35.83	7.48	33.39	10.04	30.57	7.59
Currently employed	0.09	0.28	0.07	0.25	0.08	0.27
Previously employed	0.09	0.29	0.06	0.25	0.10	0.30

Note: Column illiterate shows mean and standard deviation for women from control group who are illiterate. Column neo-literate shows for women from treatment group who were formerly illiterate. Column literate shows for women from control group who are literate. Variables own mobile, access to mobile, currently employed and previously employed take value 1 if the answer is yes and take value 0 otherwise.

Variables Number of adults, number of children and age take absolute values.

Variable income source takes value 1 if the income source is service;
takes value 2 if the income source is farming;
takes value 3 if the income source is agricultural labour;
takes value 4 if if the income source is rent;
takes value 5 if the income source is self employed;
takes value 6 if any other income source than those listed; and
takes value 0 if the respondent did not want to report the main source of income.

Table 3: Mobile own regressions

	(1)	(2)	(3)	(4)	(5)
Mobile own					
Informally literate	0.0843***	0.0644***	0.0698***	0.0614***	0.0755***
	(0.00449)	(0.00835)	(0.00819)	(0.00838)	(0.0133)
Literate	0.281***	0.217***	0.222***	0.197***	0.160***
	(0.00302)	(0.0127)	(0.0133)	(0.0111)	(0.0151)
Constant	0.193***	0.416***	0.416***	0.275***	0.459***
	(0.00217)	(0.0144)	(0.0147)	(0.0210)	(0.0299)
Controls	No	No	No	Common	Yes
Fixed effects					
Year	No	Yes	Yes	Yes	Yes
Class	No	Yes	No	No	No
State	No	Yes	No	No	No
State × class	No	No	Yes	Yes	Yes
N	102439	102439	102439	99615	31086

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Column (1): mobile own (mobile_own) regressed on dummies for informally_literate and literate keeping illiterate as base i.e. equation 1 without controls
Column (2): column (1) with year, class and state fixed effects
Column (3): column (1) with year and state_class interaction
Column (4): column (3) with controls common to all the years (job_current and edu_male)
Column (5): column (4) with all controls (household_size, schoolgoing_age_child and farmland_own).

Table 4: Mobile access regressions

	(1)	(2)	(3)	(4)	(5)
Mobile Access					
Informally literate	0.0236***	0.0351***	0.0329***	0.0252***	0.0397***
	(0.00578)	(0.00866)	(0.00943)	(0.00916)	(0.0116)
Literate	0.0761***	0.112***	0.109***	0.0832***	0.0595***
	(0.00400)	(0.0103)	(0.0109)	(0.00945)	(0.0108)
Constant	0.617***	0.603***	0.604***	0.498***	0.557***
	(0.00297)	(0.0300)	(0.0303)	(0.0375)	(0.0256)
Controls	No	No	No	Common	Yes
Fixed effects					
Year	No	Yes	Yes	Yes	Yes
Class	No	Yes	No	No	No
State	No	Yes	No	No	No
State × class	No	No	Yes	Yes	Yes
N	65705	65705	65705	63545	18692

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Column (1): mobile access (mobile_access) regressed on dummies for informally_literate and literate keeping illiterate as base i.e. equation 1 without controls
Column (2): column (1) with year, class and state fixed effects
Column (3): column (1) with year and state_class interaction
Column (4): column (3) with controls common to all the years (job_current and edu_male)
Column (5): column (4) with all controls (household_size, schoolgoing_age_child and farmland_own).

Table 5: Aware mobile money regressions

	(1)	(2)	(3)	(4)	(5)
Aware mm					
Informally literate	0.00346***	-0.00458***	-0.00393**	-0.00761***	-0.00532*
	(0.000809)	(0.00162)	(0.00187)	(0.00222)	(0.00279)
Literate	0.0657***	0.0424***	0.0433***	0.0309***	0.0282***
	(0.00112)	(0.00346)	(0.00339)	(0.00217)	(0.00288)
Constant	0.00364***	0.117***	0.120***	0.0737***	0.0902***
	(0.000331)	(0.00736)	(0.00586)	(0.0117)	(0.00884)
	(0.000323)	(0.00752)	(0.00653)	(0.0129)	(0.0109)
Controls	No	No	No	Common	Yes
Fixed effects					
Year	No	Yes	Yes	Yes	Yes
Class	No	Yes	No	No	No
State	No	Yes	No	No	No
State × class	No	No	Yes	Yes	Yes
N	102439	102439	102439	99615	31086

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Column (1): aware mobile money (aware_mm) regressed on dummies for informally_literate and literate keeping illiterate as base i.e. equation 1 without controls
Column (2): column (1) with year, class and state fixed effects
Column (3): column (1) with year and state_class interaction
Column (4): column (3) with controls common to all the years (job_current and edu_male)
Column (5): column (4) with all controls (household_size, schoolgoing_age_child and farmland_own).

Table 6: Use mobile money regressions

	(1)	(2)	(3)	(4)	(5)
Use mm					
Informally literate	-0.000830	-0.000476	-0.000607	-0.000535	-0.0429
	(0.000811)	(0.00124)	(0.00123)	(0.00124)	(0.0915)
Literate	0.00580***	-0.00115	-0.000946	-0.000663	-0.0566
	(0.000872)	(0.00108)	(0.00110)	(0.00114)	(0.0687)
Constant	0.00225***	0.0187	0.0213*	0.0278**	0.175*
	(0.000503)	(0.0120)	(0.0109)	(0.0121)	(0.0931)
Controls	No	No	No	Common	Yes
Fixed effects					
Year	No	Yes	Yes	Yes	Yes
Class	No	Yes	No	No	No
State	No	Yes	No	No	No
State × class	No	No	Yes	Yes	Yes
N	28153	28153	28153	28042	976

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Column (1): use mobile money (use_mm) regressed on dummies for informally_literate and literate keeping illiterate as base i.e. equation 1 without controls
Column (2): column (1) with year, class and state fixed effects
Column (3): column (1) with year and state_class interaction
Column (4): column (3) with controls common to all the years (job_current and edu_male)
Column (5): column (4) with all controls (household_size, schoolgoing_age_child and farmland_own).

Table 7: Bank A/C own regressions

	(1)	(2)	(3)	(4)	(5)
Bank A/C					
Informally literate	-0.00269	0.00960	0.0104	0.00701	0.00552
	(0.00517)	(0.00922)	(0.00965)	(0.0101)	(0.0132)
Literate	0.0887***	0.0703***	0.0710***	0.0621***	0.0304*
	(0.00345)	(0.00940)	(0.0112)	(0.0109)	(0.0165)
Constant	0.473***	0.419***	0.419***	0.315***	0.549***
	(0.00274)	(0.0121)	(0.0126)	(0.0179)	(0.0192)
Controls	No	No	No	Common	Yes
Fixed effects					
Year	No	Yes	Yes	Yes	Yes
Class	No	Yes	No	No	No
State	No	Yes	No	No	No
State × class	No	No	Yes	Yes	Yes
N	102439	102439	102439	99615	31086

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Column (1): bank A/C (bank_ac) regressed on dummies for informally_literate and literate keeping illiterate as base i.e. equation 1 without controls
Column (2): column (1) with year, class and state fixed effects
Column (3): column (1) with year and state_class interaction
Column (4): column (3) with controls common to all the years (job_current and edu_male)
Column (5): column (4) with all controls (household_size, schoolgoing_age_child and farmland_own).

Table 8: Bank A/C access regressions

	(1)	(2)	(3)	(4)	(5)
Bank A/C access					
Informally literate	0.000726	0.00324	0.00340*	0.00253	0.00774
	(0.00169)	(0.00196)	(0.00193)	(0.00191)	(0.00578)
Literate	0.0114***	0.00936***	0.00913***	0.00730***	0.00638
	(0.00133)	(0.00173)	(0.00169)	(0.00177)	(0.00445)
Constant	0.0137***	0.0336***	0.0341***	0.0276***	0.0818***
	(0.000879)	(0.00234)	(0.00232)	(0.00332)	(0.00943)
Controls	No	No	No	Common	Yes
Fixed effects					
Year	No	Yes	Yes	Yes	Yes
Class	No	Yes	No	No	No
State	No	Yes	No	No	No
State × class	No	No	Yes	Yes	Yes
N	49060	49060	49060	47578	11774

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Column (1): bank account access (bank_ac_access) regressed on dummies for informally_literate and literate keeping illiterate as base i.e. equation 1 without controls
Column (2): column (1) with year, class and state fixed effects
Column (3): column (1) with year and state_class interaction
Column (4): column (3) with controls common to all the years (job_current and edu_male)
Column (5): column (4) with all controls (household_size, schoolgoing_age_child and farmland_own).

Table 9: Mobile own TA+

	(1)	(2)	(3)	(4)
Mobile own				
Treatment	0.114***	0.131***	0.1000***	-0.140***
	(0.0216)	(0.0225)	(0.0285)	(0.0277)
Constant	0.435***	-0.131		0.0804
	(0.0176)	(0.125)		(0.223)
Controls	No	Yes	Yes	Yes
N	2352	2349	2349	1998

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Column (1): mobile own (mobile_own) regressed on treatment keeping only illiterate from the control sample as base i.e. equation 2 without controls

Column (2): column (1) with controls;

Column (3): column (2) using PSM (women with similar characteristics from treatment and control samples are matched);

Column (4): dependent variable regressed on treatment keeping only literate from the control sample as base i.e. equation 2 with controls

Table 10: Mobile access TA+

	(1)	(2)	(3)	(4)
Mobile access				
Treatment	0.0867***	0.0864***	0.0729	-0.0130
	(0.0208)	(0.0223)	(0.0463)	(0.0269)
Constant	0.826***	0.853***		0.940***
	(0.0179)	(0.0812)		(0.0869)
Controls	No	Yes	Yes	Yes
N	1150	1149	1149	830

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Column (1): mobile access (mobile_access) regressed on treatment keeping only illiterate from the control sample as base i.e. equation 2 without controls

Column (2): column (1) with controls;

Column (3): column (2) using PSM (women with similar characteristics from treatment and control samples are matched);

Column (4): dependent variable regressed on treatment keeping only literate from the control sample as base i.e. equation 2 with controls

Table 11: Bank account TA+

	(1)	(2)	(3)	(4)
Mobile access				
Treatment	-0.00766	0.0512***	0.0506*	-0.0529**
	(0.0175)	(0.0181)	(0.0300)	(0.0216)
Constant	0.802***	-0.161		-0.108
	(0.0141)	(0.121)		(0.132)
Controls	No	Yes	Yes	Yes
N	2352	2348	2348	1998

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Note: Column (1): bank account (bank_ac) regressed on treatment keeping only illiterate from the control sample as base i.e. equation 2 without controls
Column (2): column (1) with controls;
Column (3): column (2) using PSM (women with similar characteristics from treatment and control samples are matched);
Column (4): dependent variable regressed on treatment keeping only literate from the control sample as base i.e. equation 2 with controls

Table 12: Pairwise t-tests for dependent variables by literacy levels

	(1)	(2)	(3)	(4)	(5)	(6)
	mobile_own	mobile_access	bank_ac	bank_ac_access	use_mm	aware_mm
1 vs 2	0.248***	0.593***	0.482***	0.0149***	0.00201**	0.00478***
	(0.00221)	(0.00261)	(0.00229)	(0.000975)	(0.000953)	(0.00115)
1 vs 3	0.379***	0.607***	0.506***	0.0164***	0.00171	0.00996***
	(0.00320)	(0.00416)	(0.00332)	(0.00145)	(0.00136)	(0.00166)
2 vs 3	0.625***	0.658***	0.627***	0.0293***	0.0160***	0.103***
	(0.00140)	(0.00235)	(0.00146)	(0.000731)	(0.000579)	(0.000730)
N						

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

1=illiterate; 2= informally literate and 3=literate