

A Tough Call: Understanding the Impact of Mobile Technology on Women’s Work, Gender Gaps, Social Norms, and Misinformation

Giorgia Barboni University of Warwick	Anwesha Bhattacharya Harvard University	Erica Field Duke University
Charity Troyer Moore Yale University	Rohini Pande Yale University	Natalia Rigol Harvard University
Simone Schaner University of Southern California	Aruj Shukla University of Southern California	

May 17, 2023

Contents

1	Motivation	1
2	Study Context	2
2.1	The SKY Program	2
2.2	Study Population	2
3	Empirical Approach	3
3.1	Sampling	3
3.2	Empirical Approach: Local Randomization	4
3.3	Empirical Approach: Continuity-Based Methods	5
3.4	Estimating Local Average Treatment Effects	5
3.5	Aggregating within Families of Outcomes	6
3.6	Identification Assumptions	6
4	Data	6
4.1	Overview of Data Sources	6
4.2	Primary Data Collection	7

5	Research Questions and Hypotheses	8
6	Heterogeneity	15

1 Motivation

In lower-income settings like India, the rapid spread of mobile phones has connected poor individuals to vital information, markets and services. Previous research shows that mobile phones improve individual well-being and promote productivity by enhancing market functioning and enabling value-added services such as mobile money, information provision, and reminders.¹ However, a second generation of research is needed to understand the impacts of smartphones, which are rapidly giving individuals in the developing world their first access to the Internet and social media platforms.

Studying smartphone-based mobile Internet is crucial to understanding life in the developing world: In India, 24 percent of citizens had a smartphone in 2017, with ownership expected to double to 48 percent by 2025 (GSMA, 2018). With nearly a billion projected users, smartphones and mobile technology are poised to reshape India’s society and economy. This period of rapid expansion offers a unique opportunity to learn about these effects and their distributional consequences.

We study a statewide program targeted to women in rural households in one of India’s poorest and more conflict-prone states. In 2018, the Government of Chhattisgarh (GoCG) connected millions of these women to Internet-enabled smartphones. Under the *Sanchaar Kranti Yojana* (SKY) program, GoCG distributed over 2 million smartphones to rural Chhattisgarhi households. The aim was to bring women and their families online, most for the first time: the government ensured all program villages had LTE (i.e., 4G) coverage, and all phones received 1 GB of free data each month for the first six months.

SKY facilitated first-time mobile Internet access, which arrived en masse to remote rural communities. Such a significant change is likely to impact a wide variety of domains touching on economic, social and political life. We explore these possibilities in turn below, outlining the ways in which we anticipate SKY affected rural communities, the research questions they raise, and the data we will collect to answer these questions.

Prior to outlining these questions, we explain the SKY program and our study population, our approach to causal analysis of its impacts, and the data and empirical approaches we will use to answer these questions.

¹On market impacts, see Jensen (2007); Aker (2010); Aker and Fafchamps (2014); Tack and Aker (2014); Jensen and Miller (2018); Gupta et al. (2019). On mobile money, see Jack et al. (2013); Blumenstock et al. (2015, 2018); Suri and Jack (2016). On information/ reminders, Aker et al. (2012); Cunha et al. (2017); Dammert et al. (2015); Cadena and Schoar (2011); Dammert et al. (2014); Kast et al. (2013); Karlan et al. (2012, 2016); Zurovac et al. (2011).

2 Study Context

2.1 The SKY Program

A key challenge—especially when access to mobile technology shifts at scale—is that such access is fundamentally endogenous. The Sanchar Kranti Yojana (SKY) program provides a unique opportunity to assess the causal impact of first-time mobile Internet on rural populations. In addition to its female-friendly design, SKY distribution was motivated and promoted in the context of an upcoming statewide election, enjoying high-visibility promotion by senior state leaders, including the incumbent Chief Minister. As a result, the program was well-implemented and reached the majority of citizens in its target area.²

SKY was implemented at the Gram Panchayat (village cluster, or GP) level, with program locations determined based on a GP population threshold, which allows us to use a regression discontinuity design to study the impact of the program.³ After population-eligible villages were enumerated, the government listed areas that did not have LTE (4G) network coverage and subcontracted tower construction to the telecommunications company that provided SKY SIM cards.

Our analysis of Indian census data combined with program administrative data confirms strong adherence to program rules. Just 2 percent of population-ineligible villages received SKY, while 74 percent of eligible villages benefited from the program. When distribution occurred, an average of 0.8 phones were distributed per household, close to the program’s goal.

2.2 Study Population

Home to 25.5 million people, 40 percent of whom live in poverty, Chhattisgarh is India’s poorest state (RBI, 2019). As of the 2011 Census, 39 percent of women and 19 percent of men were illiterate. There is a great deal of spatial variation within the state, however—the central, more urban areas are relatively well off, while individuals living in the more remote rural areas are substantially poorer and often belong to historically marginalized demographic groups, called scheduled castes and tribes. Some of the more remote areas also suffer from

²Phone distribution stopped in October 2018 to comply with the Election Commission of India’s rules that limit the implementation of new welfare programs in advance of elections, though tower construction continued.

³To be eligible, a GP had to have at least one village with a population of 1,000 or more per the 2011 Indian census.

high rates of political extremist activity, perpetrated by a group called the Naxals.⁴

According to the 2019-21 National Family Health Survey, 41 percent of Chhattisgarhi women aged 15-49 have a mobile phone that they themselves use; this is well below the 54 percent nationwide average, and ranks Chhattisgarh second-to-last among all Indian states (International Institute for Population Sciences (IIPS) and ICF, 2022). Our own qualitative research in Chhattisgarh suggests that gender norms are an important barrier for women engaging with phones: Many families see phones as an unwelcome distraction that exposes women to ideas and individuals that threaten women’s purity (Barboni et al., 2018).

3 Empirical Approach

3.1 Sampling

SKY eligibility was based on a population threshold using the 2011 Indian Census. Mobile coverage was extended to covered communities, and smartphones were distributed to one adult female per household in GPs with a population of at least 1,000 in their largest village. GPs just under this population threshold were ineligible. To maximize power, our data collection strategy will focus on GPs closest to the discontinuity. We determined our sample as follows:

- First, we shortlisted the 13 districts that had the highest rates of SKY implementation, excluding the capital district of Raipur, where we had run a randomized controlled trial in many SKY eligible GPs.
- Then, using 2011 Indian census data and local randomization approaches described in Cattaneo et al. (2023), we identified a window around the population discontinuity within which one might possibly assume that treatment is as good as randomly assigned. The local randomization algorithm works as follows: we begin by limiting the sample of GPs to a window of size 10 around the discontinuity and test for balance using an F-test of the hypothesis that GP characteristics have no effect on the likelihood of being chosen as an SKY GP.⁵ We progressively widen this window, as long

⁴The Naxals descend from the Communist Party of India and attempted to derail the 2018 and 2019 elections by calling for a boycott, holding demonstrations encouraging individuals to stay away from the polls, and detonating several improvised explosive devices at polling places.

⁵We use the following GP-level variables to check for balance within the local randomization algorithm: average household size, fraction population female, fraction population scheduled caste, fraction population scheduled tribe, a dummy to identify GPs with a village unconnected to a tarmac road, area and hectares per household, number of primary schools per 1000 households, number of middle schools per 1000 households, area sown per household, a dummy to identify GPs with mobile phone coverage, a dummy to identify GPs

as the p-value from the F test is equal to 0.15 or greater. All p-values are calculated using randomization inference. We find that when dropping GPs that have just one village, the local randomization approach admits GPs where the population of the largest village is 1000 ± 99 . This yields a sample of 687 GPs.

- Depending on budgetary considerations, we may need to drop two of the most remote districts within our shortlist. Doing so leaves the local randomization threshold unchanged, but reduces the number of candidate GPs to 624.
- We expect the realized sample size will be slightly smaller than the numbers listed above, since some GPs may not be surveyable (e.g. due to safety issues due to left wing extremism or challenges in obtaining permission from local leaders).

3.2 Empirical Approach: Local Randomization

To maximize power, our main analysis will use the local randomization approach. This means our analysis will use OLS estimates leveraging the full set of surveyed GPs and including district fixed effects (which we anticipate will further improve power). Our main analysis will focus on intent to treat (ITT) effects. Formally, the regression specification is:

$$Y_{ig} = \beta_0 + \beta_1 1\{pop_g \geq 1000\} + \delta_g + \varepsilon_{ig} \quad (1)$$

where Y_{ig} is the outcome of interest in GP g for individual i , pop_g is the 2011 Census population of the largest village in GP g , δ_g is a vector of district fixed effects, and ε_{ig} is an error term.

Since many of our research questions center about gender disaggregated effects and/or impacts on gender gaps, we will also estimate heterogeneous treatment effects by gender using the following specification:

$$Y_{ig} = \lambda_0 + \lambda_1 1\{pop_g \geq 1000\} + \lambda_2 1\{pop_g \geq 1000\} \times male_{ig} + \lambda_3 male_{ig} + \delta_g + \varepsilon_{ig} \quad (2)$$

where $male_{ig}$ is a dummy variable identifying male respondents. For all specifications, we will cluster standard errors at the GP level whenever there are multiple observations per GP. Single GP specifications will use heteroskedasticity robust standard errors.

with a post office, a dummy to identify GPs with bus service, a dummy to identify GPs with self-help groups, a dummy to identify GPs with a bank or cooperative, a dummy to identify GPs with a fair price shop, and a dummy to identify GPs in districts affected by the left wing extremism. This final variable comes from a list of districts shared with us by the state.

We conduct balance checks to ensure that the control and treatment villages are not significantly different and results are shown in Table 1.

3.3 Empirical Approach: Continuity-Based Methods

We will also conduct robustness checks using a continuity-based approach. The running variable for the regression discontinuity is the population of the largest village in the GP. Then the intent to treat effect is:

$$\tau_{ITT} = \lim_{pop \downarrow 1000} E[Y_{ig} | pop_g = pop] - \lim_{pop \uparrow 1000} E[Y_{ig} | pop_v = pop] \quad (3)$$

where Y_{iv} is the outcome of interest in village v for individual i . We use a bandwidth of ± 100 to minimize standard errors and to ensure all sampled villages get a positive weight. Results from the balance checks using this method are in Table 2. This table also verifies that results are very similar but less well-powered when using mean squared error optimal bandwidths per (Calonico et al., 2014).

3.4 Estimating Local Average Treatment Effects

Since not all eligible communities received SKY (the program was paused in advance of the 2018 state election and not taken up by the new administration), we also anticipate conducting analysis using a fuzzy RD design. In the local randomization approach, we will use the eligibility variable as an instrumental variable (IV) for actual distribution of SKY phones to estimate local average treatment effects (LATE).

Similar to τ_{ITT} , we can estimate LATE using a continuity-based approach to check for robustness of our results.

$$\tau_{LATE} = \frac{\tau_{ITT}}{\tau_{FS}} = \frac{\lim_{pop \downarrow 1000} E[Y_g | pop_v = pop] - \lim_{pop \uparrow 1000} E[Y_g | pop_v = pop]}{\lim_{pop \downarrow 1000} E[D_g | pop_g = pop] - \lim_{pop \uparrow 1000} E[D_g | pop_g = pop]} \quad (4)$$

where $D_v = 1$ if SKY distribution took place in village v and $D_v = 0$ otherwise.

We may not include these specifications in our final analysis. If we do, we will largely use this as supplementary analysis to facilitate thinking about magnitudes, since tower construction affected some communities where distribution never took place.

3.5 Aggregating within Families of Outcomes

Whenever relevant we will group related variables into families of outcomes and focus results on an aggregated standardized index that summarizes family-wide effects, while reporting impacts on individual index components in supplementary results.

3.6 Identification Assumptions

The central assumption underpinning our causal analysis is that SKY receipt is “as good as” randomly assigned within our chosen window of ± 99 around the eligibility threshold (or, in the continuity-based approach, at the threshold). We verify that this assumption holds using pre-SKY village-level characteristics from the 2011 census data (Table 1, 2). We also use tests proposed by McCrary (2008); Cattaneo et al. (2019) and cannot reject the null of no bunching of the running variable (population) around the cutoff, confirming there was no manipulation around the eligibility cutoff. Finally, a review of major national and state programs that utilize population thresholds for implementation has not identified major issues regarding other programs using the same eligibility cutoff. In brief, data suggest our identification strategy will generate valid causal estimates.

4 Data

4.1 Overview of Data Sources

We will utilize administrative, secondary, and primary data for our intended analyses. Program data on SKY implementation include information on whether a village was selected for SKY, if and when phone distribution took place, and how many phones were distributed. We also have information on whether an area was slated to receive a new LTE tower and when the tower came online.⁶

The bulk of our analysis will be enabled by primary data collection—including individual and village-informed respondent surveys—that will be undertaken beginning in early 2023. We will combine data from these surveys with program administrative data to compare SKY impacts, with a focus on locations just above and below the program eligibility threshold.

Our analysis will also leverage data from a 2011 Indian Census, administrative data on voter turnout and electoral outcomes, data on political candidate criminal backgrounds from

⁶We have digitized and merged these data with the 2011 Indian Census to verify program rules were followed.

the Association for Democratic Reforms, administrative data from the Mahatma Gandhi National Rural Employment Guarantee Scheme MIS, and Speedtest by Ookla Global Fixed and Mobile Network Performance Maps. We may include additional secondary data sources, as our research develops. This pre-analysis plan is focused on outcomes collected in our primary survey data (see below), since we can credibly signal our interest in studying those outcomes before the data are collected.

4.2 Primary Data Collection

Gram Panchayat Key Informant Survey: The Gram Panchayat survey will cover up to 687 rural Gram Panchayats (GPs) across up to 13 districts of Chhattisgarh. We will conduct 2 village-level key informant interviews per GP to measure village-level outcomes. We anticipate these surveys will be completed by a “town crier”, broadly familiar with economic activities in the community, and a local ASHA worker, a female government employee focused on village health and nutrition.

Individual Survey: We will interview both adult females (aged 18-45) and males (aged 18-50) to study how outcomes vary by gender, and whether norms evolve differently among men and women. We will conduct in-person surveys of 15 men and 15 women per GP and limit attention to households that have at least one age-eligible woman the relevant bits are in the primary data collection subsection.⁷

To ensure broad representativeness in the sample, we will utilize publicly available voter registration lists for every polling station in the sample GPs. These lists have key details including an individual’s name, address, age, and gender. Importantly, voter registration in India is high, with 911 million registered voters ⁸. The set of registered voters will be aggregated at the GP level. We will then draw a random sample of age-eligible adult females from this list. Trained field officers will track the sampled woman, and verify her eligibility and availability. If she is eligible, we shall enroll her in our study. If she is not eligible, the field officer will create a list of eligible women in the same household and one respondent will be randomly selected from this list. At this point, the household will be enrolled in the sample. If the household has no eligible female respondent, we will drop this household from our sample and move on to a replacement household. Once the female respondent has been selected, the officer will ask her to list down all male household members who are age-eligible. Using a similar listing process as above, a male respondent is randomly selected

⁷Whenever possible, we will interview a man and a woman from the same household.

⁸Source: <https://eci.gov.in/statistical-report/statistical-reports/>

from this list. If there is no age-eligible man available to be enrolled, we will enroll a man from a replacement household that has an age eligible woman.

5 Research Questions and Hypotheses

SKY may have affected a variety of individual, household and community-level variables in domains relevant to economic and social outcomes. To facilitate understanding, we describe our hypotheses and research questions in broad groupings below. While we anticipate exploring research questions through the frameworks outlined here, we will ultimately communicate research findings in a way that best reflects the insights the data provides.

“First Stage Effects”: Did SKY have lasting effects on access to 4G internet and use of smartphones?

While we know from administrative data that SKY boosted smart phone ownership in the short run, differences may have attenuated over time (e.g. as phones depreciate/break, or towers are built in uncovered areas). We will measure present-day access to mobile technology and the internet in two ways: First, by measuring phone ownership and use in individual surveys. In order to paint a holistic picture of phone use, we will construct indices that measure use of basic (button) phones and engagement in tasks that do not require a smartphone (making a call, picking up a call, sending an SMS etc), as well as use of smartphones and tasks that require a smartphone. Second, we will conduct mobile internet speed tests in each GP we survey to measure network strength and speed.

Understanding the durability of SKY’s impacts on phone use and network access is interesting in its own right, and relevant for informing our assessment of the program’s impact on other downstream outcomes.

We classify the remainder of our research questions into four conceptually distinct domains, enumerated below.

Research Area 1: Individual and Gendered Impacts of Mobile Internet

The distribution of smartphones and free data provision enabled first-time mobile Internet access across rural Chhattisgarh. Similar to the arrival of 2G and 3G in low-income settings, we anticipate SKY catalyzed changes in individual and household-level access to information,

networks, markets, and services. These changes likely affected decisions related to financial products and tools, jobs and migration.

Moreover, SKY’s gender-targeted distribution policy may have played an important role in closing a variety of gender gaps, both economic and social. We hypothesize that the implementation of SKY, which involved a series of enrollment camps for women in central village locations, positively signaled that it is appropriate for women to use phones. Another channel through which SKY likely operated comes through its scale-changing norms may require concentrated behavior change among large swathes of the population. On the other hand, if conservative norms depress women’s access or returns to digital technology, a marginalization trap could emerge where norms exacerbate gender gaps in technology adoption. Marginalization of women may deepen conservative gender norms, creating a negative feedback loop.

Given the wide scope for gendered defects (beyond income generation phones could, for example, correct for pre-existing gendered information gaps that lead to under-investment in children, poor utilization of health services, and limit the reach of state-sponsored social protection programs), our analysis in this research area will focus on both economic outcomes and social norms. Below, we specify which outcomes are only measured for women. When outcomes are available for both genders, we will study both overall impacts as well as gender disaggregated effects.

Research Question 1.1: How did SKY change men’s and women’s engagement with smartphones, the internet and digital gender gaps?

We expect SKY to increase phone usage – however it is not obvious if men or women will benefit more, especially given potential for backlash. Specific outcomes related to this question include engagement with phone calls, use of the internet, access to digital financial services through mobile phone applications, and use of social media applications like YouTube, Facebook and WhatsApp. We will specifically explore whether SKY closed gender gaps in these areas.

Research Question 1.2: Did SKY influence norms around women’s phone use?

We will collect detailed survey data on attitudes and norms governing women’s use of mobile phones, studying how SKY impacts both first and second order beliefs by gender.

Research Question 1.3: Did SKY impact labor force participation and economic activities?

With increased access to smartphones, we hypothesize that survey respondents will be better able to learn about job opportunities and connect with individuals and information relevant to income generation. Villagers may also access online platforms to learn a new skill useful to earning income (e.g., stitching). Primary outcomes of interest include labor market participation, employment, and income. We will also assess participation in different sectors of the labor market. To better understand mechanisms, we will also explore how survey respondents in SKY communities use phones and mobile Internet and the extent to which they use them to engage in labor markets.

Mobile Internet access may also affect migration for work: increasing it as individuals learn about relative wages and new job opportunities, or reducing it as they are able to earn an income through phone-based activities. In addition to collecting data on household members who have recently migrated, we will collect information on remittances and perception of wages in higher-return urban labor markets.

It is not obvious whether labor market changes will differentially accrue to men versus women. We will therefore study overall impacts on labor market outcomes, and how SKY affects the relative participation of men and women in different sectors of the labor market and engagement in income generation activity. We will also study how the program impacted women's time use.

Research Question 1.4: Did mobile Internet access change how individuals engage with financial markets?

With the ability to use phones to view bank account balances and undertake digital financial transactions, individuals in SKY villages might visit the bank less often, but engage more frequently with new digital payment services. To understand this, we will assess impacts on individuals' awareness, understanding and use of analog and digital financial services, and their use of mobile phones to access financial services. We will focus on both overall impacts and gender disaggregated effects to understand whether SKY affects gender gaps in financial activity.

Research Question 1.5: Did SKY affect women's networks and social connections?

Phones can change the frequency with which women communicate with friends and family. This could increase social connections and have a positive impact on digital connectivity—possibly at the cost of in-person interactions with friends. We include a survey module designed to measure women's social network size and the extent to which they interact with network connections in person and on the phone. We will also measure impacts on self-help

group membership and activity.

Research Question 1.6: Did SKY have any impact on broader female empowerment outcomes, including decision making and experiences of emotional violence?

Access to mobile Internet could directly affect indicators of female empowerment through a variety of channels. For example, media available on mobile Internet (e.g., YouTube) may present women (and men) with examples of female role models who exercise more agency than local women, affecting gender norms and views on women’s roles in the household and the economy. Moreover, if phones provide new income generating opportunities, women could gain intra-household power and seek to exercise more agency or a larger role in household decisions. On the other hand, backlash to the program might stall women’s progress on these outcomes or even result in backsliding.

To understand questions of broader female empowerment, we measure respondents’ participation in common household decisions (Jayachandran et al., 2021) as well as a Demographic and Health Survey-style module on controlling behavior.

Research Question 1.7: Did SKY affect female mobility and perceptions of safety?

SKY might facilitate increased mobility among women if carrying a phone makes it safer for them to travel long distances. On the other hand, increased access to information (say, on crime) through mobile phones might create the impression that the world outside their village is more unsafe for women, thereby limiting mobility. We measure both actual mobility and perceptions of safety in our survey to understand if any of these mechanisms are at play.

Research Question 1.8: Did SKY affect women’s broader sense of well-being and mental health?

To examine how the arrival of mobile phones and the reshaping of social connectivity can impact women’s well-being, we will administer a depression and anxiety checklist using the standardized Patient Health Questionnaire-4 (PHQ4) scale. We will also measure loneliness via a subset of questions from the UCLA loneliness scale.

Research Area 2: Community-Level Impacts of Mobile Internet

Beyond its impact on individual- and household-level outcomes, SKY may have shifted broader general equilibrium outcomes including prices and wages as information asymmetries were closed across locations.

Within this domain, we will explore the following key research questions:

Research Question 2.1: Did SKY affect the presence of local businesses and aggregate market outcomes?

Over time, given changes in individual economic decision making, broader market-level changes may occur that reshape the local economy. We will use data from village informant surveys to assess SKY’s impact on the number and type of local businesses, labor market tightness, wages, and presence of different types of financial service providers.

By closing information gaps and facilitating remote transactions, mobile Internet may affect the presence and use of financial services in rural areas. The role of mobile Internet is unclear *ex ante*, since remote services may substitute for local service points, but increased ability to transact or increases in economic activity could generate local complementarities that increase in-person transactions.

Research Question 2.2: Does SKY have any impact on the distribution of good prices or on local wage rates?

Prior evidence has shown that mobile phones facilitate price arbitrage, improving the functioning of markets for perishable goods Jensen (2007). Using data on what households pay for a small basket of consumption goods including rice, sugar, pulses, tomatoes, onions, eggplant, and green chilies, we will study SKY’s impact on moments of the price distribution for both perishables and non-perishables. We will similarly measure impacts on the distribution of wage rates.

Research Area 3: Health, Information, and Coping during COVID-19

First-time mobile Internet access arrived in communities approximately one year prior to the beginning of the Covid-19 pandemic, which had devastating economic and health consequences across India. Pandemic impacts were mediated in part through access to information to prevent infection and spread, and encourage early vaccination. Yet misinformation about Covid-19 was widespread, driven in part through inaccurate social media posts. The rise of “fake news” is significant enough to be deemed an “infodemic” of mis- and dis-information, exacerbated by the pandemic (WHO, 2020), which could result in poorer knowledge and behavior outcomes (Roozenbeek et al., 2020; Lee et al., 2020). Under this research area, we explore the extent to which first-time mobile Internet access affected access to both accurate

and inaccurate information, and then examine whether Covid-specific outcomes differ across SKY-eligible and -ineligible locations.

Research Question 3.1: Did mobile Internet access change sources and patterns of information consumption among rural men and women?

SKY may have changed how individuals processed and assessed new information and the perceived trustworthiness of different news sources. We anticipate that it increased individuals' exposure to online news sources, and had an ambiguous impact on how much people trust the news they see online. Having more phones in a village might lead to a mechanical increase in social media consumption and trust in the information shared on the internet. On the other hand, increased exposure to social media might warrant suspicion of "viral" news and decrease trust in certain social media platforms, such as WhatsApp. We will collect information on respondents' most commonly used sources of information, as well as their trust in a broad set of information sources including traditional sources (newspapers and television), social media platforms, and communication with friends and family, local leaders, and health workers. We will also study which sources they used to receive Covid-19 information to understand whether SKY shifted the composition of the Covid-19 information pipeline.

Research Question 3.2: Does mobile Internet access affect susceptibility to misinformation?

India is the world's biggest consumer of Covid-19-related misinformation (Al-Zaman, 2022). Such "fake news" could also have a disproportionate effect on women, given the pre-existing digital gender gap, women's lower education levels, and social isolation (Heise et al., 2019). To assess susceptibility to misinformation, we will collect data on respondents' knowledge of and belief in common Covid-19 rumors circulated on social media in the study area.

Research Question 3.3: Does mobile Internet access correlate with Covid-related knowledge and preventive behaviors?

The consequences of both accurate and inaccurate information available through mobile Internet include health knowledge and health-seeking behaviors. To assess this, our survey will ask respondents about knowledge of and belief in useful facts about Covid-19 and record the vaccination status of all members of surveyed households.

Research Question 3.4: Does mobile Internet access impact village-level incidence of

Covid-19 and Covid-related impacts?

The net effect of SKY on Covid outcomes is unclear: while phones may have facilitated access to useful information and protective technology like vaccines, misinformation may have worsened infection and death rates.

To study net impacts on Covid-related outcomes (which we expect will also reflect accumulated adherence to preventative behaviors like social distancing and masking), we will collect data on the number of positive tests and Covid-19 deaths from village level informants. We will also collect data on Covid illness episodes and deaths from households, alongside healthcare utilization and expenditure.

We will also collect data on non-Covid illness episodes from households. We have added checks to this module to assess whether illnesses are “likely undiagnosed” Covid cases (based on, e.g., loss of taste/smell and trouble breathing). If our analysis suggests that this module is capturing many Covid cases, we may include data from this module in our analysis.

Research Question 3.5: Has mobile Internet access affected children’s human capital investment and education outcomes?

In light of concerns around school dropout rates and Covid-19, we will assess whether SKY impacted school enrollment and dropout rates differ across SKY/non-SKY villages. We will also measure the extent to which children used smart phones for school work following the onset of Covid.

Research Area 4: Political Impacts of Mobile Internet

Research Question 4.1: Has mobile Internet access affected voter turnout and election outcomes?

Prior to drafting this PAP, we analyzed administrative data on voter turnout and found preliminary evidence that SKY is associated with lower rates of turnout during the 2018 state legislative assembly and 2019 general elections. Looking ahead, we aim to deepen our analysis of electoral outcomes using secondary data. For example, we will explore whether SKY facilitated village-level coordination, as evidenced by the share of votes captured by a single candidate. We will also explore impacts on support for the incumbent political party at the time of SKY distribution, support for female candidates, support for candidates for marginalized groups, and support for candidates with criminal records. We also plan on exploring heterogeneous treatment effects with respect to pre-SKY political alignment and

community remoteness.

Research Question 4.2: Did SKY impact political engagement and knowledge?

Reduced turn out may reflect less interest/engagement in political matters. To test this hypothesis, our survey measures both political knowledge (specifically respondents' knowledge of key officeholders such as President and Prime Minister of India, as well as local representatives) and self-reported political engagement.

Research Question 4.3: Does mobile Internet have any impact on citizens' trust in institutions and affective polarization?

Another concern related to the unintended negative effects of exposure to harmful misinformation, often arriving through mobile Internet, is that it could increase vulnerability to scams and fraud, and cause them to be disillusioned with the state (Guriev et al., 2019). We will study impacts on a set of interpersonal and group trust questions derived from the Afrobarometer questionnaires. Different levels of trust could affect how individuals engage with political institutions and how pivotal they feel their voice to be. We measure this by asking respondents if they participate in protests or demonstrations, or engage with government officials to resolve public goods provisioning issues. In addition, we measure the magnitude of affective polarization through a series of social distance questions.

Research Question 4.4: Does mobile Internet access affect collective action and protests?

Communication and fast dissemination of information are major determinants of collective action (Pierskalla and Hollenbach, 2013; Shapiro and Weidmann, 2015). Access to mobile phones can facilitate coordination among protesting groups and hence we may see an increase in demonstrations and protests, or participation in other types of collective action. On the other hand, access to phones can drive down violence by bridging the gap between complainants and public service providers.

6 Heterogeneity

To better understand underlying mechanisms, we anticipate assessing heterogeneous impacts of the SKY program along the following dimensions:

I. Gender and gender norms

- A. at the individual level, we will disaggregate treatment effects by gender and study gender gaps as described earlier in this document
- B. at the GP or village level, we may examine differences in treatment effects by the sex ratio and – if we are able to obtain the data for our entire sample, the female labor force participation rate from the 2011 census⁹

II. Education

- A. self-reported education levels from the primary survey
- B. if we are able to obtain the data for our entire sample, we may also examine heterogeneity in local education levels as proxied by literacy levels in the 2011 census

III. Household social group

- A. whether the household is part of a marginalized social group
- B. proportion of population belonging to marginalized social groups per 2011 census

IV. Remoteness

- A. road infrastructure
- B. distance to urban centers

V. For the political outcomes we can look at effect by incumbent party leaders

- A. party of the incumbent leader (BJP vs. Indian National Congress vs. others)
- B. vote share of BJP in the previous election
- C. years of incumbency

We expect that our initial analysis may generate follow-on hypotheses that cannot be adequately addressed via the tests specified above. We will disclose in our research outputs which tests were pre-specified and which were not.

⁹At the time of drafting this PAP, we only had access to village-level census data from the village and town directory file from the district census handbook, which lacks information on labor force participation and literacy.

References

- Aker, J. C. (2010). Information from markets near and far: Mobile phones and agricultural markets in niger. *American Economic Journal: Applied Economics* 2(3), 46–59.
- Aker, J. C. and M. Fafchamps (2014). Mobile phone coverage and producer markets: Evidence from west africa. *World Bank Economic Review* 29(2), 262–292.
- Aker, J. C., C. Ksoll, and T. J. Lybbert (2012). Can mobile phones improve learning? evidence from a field experiment in niger. *American Economic Journal: Applied Economics* 4(4), 94–120.
- Al-Zaman, M. S. (2022). Prevalence and source analysis of covid-19 misinformation in 138 countries. *IFLA Journal* 48(1), 189–204.
- Barboni, G., E. Field, R. Pande, N. Rigol, S. Schaner, and C. Troyer Moore (2018). A tough call: Understanding barriers to and impacts of women’s mobile phone adoption in india.
- Blumenstock, J., M. Callen, and T. Ghani (2018, October). Why do defaults affect behavior? experimental evidence from afghanistan. *American Economic Review* 108(10), 2868–2901.
- Blumenstock, J. E., M. Callen, T. Ghani, and L. Koepke (2015). Promises and pitfalls of mobile money in afghanistan: evidence from a randomized control trial. In *Proceedings of the Seventh International Conference on Information and Communication Technologies and Development*.
- Cadena, X. and A. Schoar (2011). Remembering to pay? reminders vs. financial incentives for loan payments. Technical Report 17020, National Bureau of Economic Research.
- Calonico, S., M. D. Cattaneo, and R. Titiunik (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica* 82(6), 2295–2326.
- Cattaneo, M. D., N. Idrobo, and R. Titiunik (2023). A practical introduction to regression discontinuity designs: Extensions. *arXiv preprint arXiv:2301.08958*.
- Cattaneo, M. D., R. Titiunik, and G. Vazquez-Bare (2019). Power calculations for regression-discontinuity designs. *The Stata Journal* 19(1), 210–245.
- Cunha, N., G. Lichand, R. Madeira, and E. Bettinger (2017). What is it about communicating with parents? Technical report.

- Dammert, A. C., J. Galdo, and V. Galdo (2015). Integrating mobile phone technologies into labor-market intermediation: a multi-treatment experimental design. *IZA Journal of Labor & Development* 4(1), 1–27.
- Dammert, A. C., J. C. Galdo, and V. Galdo (2014). Preventing dengue through mobile phones: evidence from a field experiment in peru. *Journal of Health Economics* 35, 147–161.
- GSMA (2018). Global Mobile Trends: What’s Driving the Mobile Industry? Technical report.
- Gupta, A., J. Ponticelli, and A. Tesei (2019). Technology Adoption and Access to Credit Via Mobile Phones. *Available at SSRN*.
- Guriev, S., N. Melnikov, and E. Zhuravskaya (2019, 01). 3g internet and confidence in government. *SSRN Electronic Journal*.
- Heise, L., M. E. Greene, N. Opper, M. Stavropoulou, C. Harper, M. Nascimento, D. Zewdie, G. L. Darmstadt, M. E. Greene, S. Hawkes, L. Heise, S. Henry, J. Heymann, J. Klugman, R. Levine, A. Raj, and G. Rao Gupta (2019). Gender inequality and restrictive gender norms: framing the challenges to health. *The Lancet* 393(10189), 2440–2454.
- International Institute for Population Sciences (IIPS) and ICF (2022). National Family Health Survey (NFHS-5), 2019-21: India Report. Technical report.
- Jack, W., A. Ray, and T. Suri (2013). Transaction networks: Evidence from mobile money in kenya. *American Economic Review: Papers and Proceedings* 103(3), 356–361.
- Jayachandran, S., M. Biradavolu, and J. Cooper (2021, March). Using machine learning and qualitative interviews to design a five-question women’s agency index. Working Paper 28626, National Bureau of Economic Research.
- Jensen, R. (2007). The digital provide: Information (technology), market performance, and welfare in the south indian fisheries sector. *Quarterly Journal of Economics* 122, 879–924.
- Jensen, R. and N. Miller (2018). Market integration, demand, and the growth of firms: Evidence from a natural experiment in india. *American Economic Review*. Forthcoming.
- Karlan, D., M. McConnell, S. Mullainathan, and J. Zinman (2016). Getting to the top of mind: How reminders increase saving. *Management Science* 62(12), 3393–3411.

- Karlan, D., M. Morten, and J. Zinman (2012). A personal touch: Text messaging for loan repayment. Technical Report 17952, National Bureau of Economic Research.
- Kast, F., S. Meier, and D. Pomeranz (2013). Under-Savers Anonymous: Evidence on Self-Help Groups and Peer Pressure as a Savings Commitment Device. NBER Working Paper No. 18417.
- Lee, J. J., K.-A. Kang, M. P. Wang, S. Z. Zhao, J. Y. H. Wong, S. O'Connor, S. C. Yang, and S. Shin (2020, Nov). Associations between covid-19 misinformation exposure and belief with covid-19 knowledge and preventive behaviors: Cross-sectional online study. *J Med Internet Res* 22(11), e22205.
- McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics* 142(2), 698–714. The regression discontinuity design: Theory and applications.
- Pierskalla, J. and F. Hollenbach (2013, 05). Technology and collective action: The effect of cell phone coverage on political violence in africa. *American Political Science Review* 107, In Press.
- RBI (2019). Handbook of Statistics on Indian States 2018-19. Technical report, Reserve Bank of India.
- Roozenbeek, J., C. R. Schneider, S. Dryhurst, J. Kerr, A. L. J. Freeman, G. Recchia, A. M. van der Bles, and S. van der Linden (2020). Susceptibility to misinformation about covid-19 around the world. *Royal Society Open Science* 7(10), 201199.
- Shapiro, J. N. and N. B. Weidmann (2015). Is the phone mightier than the sword? cellphones and insurgent violence in iraq. *International Organization* 69(2), 247–274.
- Suri, T. and W. Jack (2016). The long-run poverty and gender impacts of mobile money. *Science* 354(6317), 1288–1292.
- Tack, J. and J. C. Aker (2014). Information, mobile telephony, and traders' search behavior in niger. *American Journal of Agricultural Economics* 96(5), 1439–1454.
- WHO (2020, September). Managing the covid-19 infodemic: Promoting healthy behaviours and mitigating the harm from misinformation and disinformation.
- Zurovac, D., R. K. Sudoi, W. S. Akhwale, M. Ndiritu, D. H. Hamer, A. K. Rowe, and R. W. Snow (2011). The effect of mobile phone text message reminders on kenyan health workers

adherence to malaria treatment guidelines: a cluster randomised trial. *Lancet* 378(9793), 795–803.

Tables and Figures

Table 1: Balance on Predetermined Characteristics from 2011 Census – Differences in Means (OLS)

	(1) Mean (Ineligible GPs)	(2) Difference (Eligible GPs)	(3) N
Average Household Size	4.452 [0.561]	-0.044 (0.042)	687
Fraction Female	0.502 [0.013]	-0.001 (0.001)	687
Fraction Scheduled Caste	0.166 [0.195]	0.007 (0.015)	687
Fraction Scheduled Tribe	0.318 [0.266]	-0.008 (0.020)	687
Has Village Unconnected to Tarmac Road	0.525 [0.500]	-0.023 (0.039)	687
Area in Hectares Per Household	1.921 [1.012]	-0.017 (0.088)	687
Number Primary Schools Per 1,000 Households	7.264 [2.493]	-0.438** (0.184)	687
Number Middle Schools Per 1,000 Households	2.998 [2.109]	-0.065 (0.149)	687
Area Sown (Hectares) Per Household	1.173 [0.426]	-0.073** (0.030)	687
Has Mobile Coverage	0.811 [0.392]	-0.033 (0.032)	687
Has Post Office	0.108 [0.311]	0.028 (0.026)	687
Has Bus Service	0.583 [0.494]	0.001 (0.038)	687
Has Self Help Groups	0.949 [0.221]	0.005 (0.017)	687
Has Bank or Co-Op	0.042 [0.200]	-0.002 (0.015)	687
Has Fair Price Shop	0.748 [0.435]	0.048 (0.032)	687
Left Wing Extremism Affected	0.429 [0.496]	-0.035 (0.038)	687

All data from 2011 Indian Census. Standard deviations in square brackets, heteroskedasticity robust standard errors in parentheses. Column 1 reports the mean and standard deviation of the variable below the SKY eligibility threshold. Column 2 reports differences in outcomes above the threshold, with heteroskedasticity robust standard errors in parentheses. Column 3 reports the total sample size.

Table 2: Balance on Predetermined Characteristics from 2011 Census – First Order Polynomial Fits

	(1)	(2)	(3)	(4)	(5)	(6)
	Bandwidth Set to 100			Optimal Bandwidths		
	Mean (Ineligible GPs)	Difference (Eligible GPs)	N	Mean (Ineligible GPs)	Difference (Eligible GPs)	N
Average Household Size	4.452 [0.561]	-0.114 (0.132)	687	4.393 [0.579]	-0.109 (0.213)	183
Fraction Female	0.502 [0.013]	-0.003 (0.004)	687	0.501 [0.014]	-0.004 (0.005)	266
Fraction Scheduled Caste	0.166 [0.195]	0.021 (0.044)	687	0.158 [0.174]	0.091 (0.074)	155
Fraction Scheduled Tribe	0.318 [0.266]	-0.059 (0.061)	687	0.325 [0.255]	-0.136 (0.096)	162
Has Village Unconnected to Tarmac Road	0.525 [0.500]	-0.007 (0.119)	687	0.583 [0.495]	0.190 (0.176)	198
Area in Hectares Per Household	1.921 [1.012]	0.211 (0.256)	687	1.808 [0.862]	-0.211 (0.526)	155
Number Primary Schools Per 1,000 Households	7.264 [2.493]	-0.395 (0.570)	687	7.319 [2.563]	-1.061 (0.786)	190
Number Middle Schools Per 1,000 Households	2.998 [2.109]	0.459 (0.455)	687	3.061 [1.982]	0.852 (0.726)	183
Area Sown (Hectares) Per Household	1.173 [0.426]	0.042 (0.105)	687	1.139 [0.402]	0.085 (0.146)	257
Has Mobile Coverage	0.811 [0.392]	0.081 (0.106)	687	0.764 [0.426]	0.171 (0.149)	237
Has Post Office	0.108 [0.311]	-0.100 (0.082)	687	0.133 [0.341]	-0.083 (0.123)	198
Has Bus Service	0.583 [0.494]	-0.012 (0.121)	687	0.514 [0.502]	-0.148 (0.167)	237
Has Self Help Groups	0.949 [0.221]	-0.093 (0.062)	687	0.969 [0.174]	-0.137 (0.087)	213
Has Bank or Co-Op	0.042 [0.200]	-0.033 (0.047)	687	0.034 [0.183]	-0.134* (0.080)	151
Has Fair Price Shop	0.748 [0.435]	-0.024 (0.101)	687	0.752 [0.434]	-0.016 (0.161)	213
Left Wing Extremism Affected	0.429 [0.496]	0.000 (0.115)	687	0.390 [0.489]	-0.011 (0.153)	257

All data from 2011 Indian Census. Standard deviations in square brackets, heteroskedasticity robust bias-corrected standard errors in parentheses. The first three columns report results where the RD bandwidth is set to 100, to ensure all sampled GPs receive positive weight. Column 1 reports the mean and standard deviation of the variable below the SKY eligibility threshold. Column 2 reports robust bias-corrected regression discontinuity estimates of differences in outcomes above the threshold, with standard errors in parentheses. Column 3 reports the total sample size. The final three columns report results with optimal bandwidths. Column 4 reports the mean/standard deviation of the outcome in GPs below the threshold that receive positive weight, column 5 reports robust bias-corrected RD point estimates and standard errors, and column 6 reports the total number of GPs in the sample receiving positive weight. All regressions use a first order polynomial and a triangular kernel. *, **, and *** indicate statistical significance at the 10, 5, and 1 percent levels respectively.