Pre Analysis Plan

1. Experimental Design and Treatment Assignment

Randomization occurred at the village-unit level. A village-unit consists of one or more geographically proximate villages grouped for sampling purposes. Henceforth, “village” refers to the randomized village-unit cluster.

Let $Z_v \in \{0,1,2\}$ denote randomized assignment:

- $Z_v = 0$: Control
- $Z_v = 1$: Treatment arm T1 (smartphones + internet)
- $Z_v = 2$: Treatment arm T2 (smartphones + internet + curated content)

We also define:

- $T_v = 1\{Z_v \in \{1,2\}\}$: Any treatment vs control.

The primary estimands are the **intent-to-treat (ITT)** effect of assignment.

---

2. Outcomes

2.1 Primary Outcome

The primary outcome is a composite **Women’s Empowerment Index** constructed from pre-specified domain indices.

All items will be coded so that higher values indicate greater empowerment.

---

3. Construction of Indices

3.1 Rescaling

All ordinal items are coded so that higher values uniformly indicate higher empowerment. Negatively framed items are reverse coded prior to aggregation.

Each item is linearly rescaled to the $[0, 1]$ interval:

$$x^* = \frac{x - \min(x)}{\max(x) - \min(x)}.$$

Within each domain, rescaled items are averaged with equal weights.

3.2 Domain-Specific Indices

(i) Financial Inclusion

Items will include measures of formal financial access (account ownership), digital financial adoption (UPI and mobile banking use), and financial autonomy (independent deposit and withdrawal decisions). Components will be rescaled to [0,1] and averaged.

(ii) Social Connectedness

Items will include both the frequency of long-distance social contact and the intensity of interaction with close social ties.

Components will be rescaled to [0,1] and averaged.

(iii) Health and Hygiene

Items will include three related dimensions of reproductive and menstrual health knowledge and practice. It will capture women’s awareness of feminine hygiene products, their menstrual protection behavior, and their awareness of a range of contraceptive methods. The awareness measures will be rescaled to lie between 0 and 1, while menstrual protection will be coded as a binary indicator distinguishing improved methods from less effective or no protection, with non-applicable responses treated as missing.

Women who report menopause or hysterectomy will be excluded from this domain. The index will be constructed as the average of the available components, and will range in [0,1].

(iv) Gender Attitudes

Items will capture respondents' views on women's roles, rights, and opportunities within the household and society. It will include measures of agreement with traditional gender role norms, such as the belief that men should be the primary earners and that boys' education is more important than girls', as well as support for women's rights to express their opinions and to work outside the home. All items will be coded so that higher values reflect more gender-equitable attitudes, and the index will be constructed as the average of the standardized components. Components will be rescaled to [0,1].

(v) Attitude towards Rejection of Intimate Partner Violence (IPV)

Attitudes toward domestic violence will be measured using three scenarios that commonly reflect normative justifications for wife-beating: whether a husband is justified in beating his wife if she argues with him, refuses to have sexual relations, or shows disrespect toward her in-laws. Responses will be coded so that higher values reflect stronger rejection of violence against women, standardized, and averaged to construct the index in [0,1].

(vi) World Bank's Empowerment Measure

Components such as decision making at the household, self-efficacy or confidence, and sharing of housework, will be combined into an average index in [0,1].

---

4. Standardization

Each domain index will be standardized using control-group mean and standard deviation(z-score):

$$Z_{id} = \frac{X_{id} - \mu_{d,control}}{\sigma_{d,control}}.$$

---

5. Composite Empowerment Index (Primary Outcome)

The composite empowerment index is defined as the simple average of standardized domain indices:

$$Women\ Empowerment_i = \frac{1}{D}\sum_{d} Z_{id}.$$

The composite index will be standardized using the control-group mean and standard deviation (z-score).

---

6. Secondary Index Construction

As a robustness exercise, we will construct a PCA-based composite index using weights estimated from the control group only.

---

7. Estimation Strategy

7.1 ITT Specification (ANCOVA)

This specification will estimate intent-to-treat (ITT) effects using an ANCOVA framework. For each outcome Y, including the composite empowerment index and domain-level indices, we will estimate:

$$Y_{iv,1} = \alpha + \beta\, Treatment_v + \rho\, Y_{iv,0} + \gamma' X_{iv,0} + \varepsilon_{iv},$$

where $Y_{iv,1}$ denotes the endline outcome for individual i in village v, $Y_{iv,0}$ denotes the corresponding baseline value, and $X_{iv,0}$ is a vector of pre-specified baseline covariates. The coefficient β captures the ITT effect of treatment assignment.

Standard errors will be clustered at the village level, the unit of randomization.

Baseline control variables will be selected using a pre-specified two-step approach. First, we will include covariates motivated by theoretical relevance to empowerment and financial decision-making, measured prior to treatment assignment. These will include demographic and socioeconomic characteristics such as age, years since marriage, age gap between spouses, household size, number of sons and daughters, religion, caste category, relationship to household head, occupation category, land ownership, SHG membership, education (respondent and husband), English proficiency, earnings, and baseline measures of preferences and behavioral traits (risk aversion, altruism, inequality aversion, impatience, present bias, loss aversion, and trust).

Second, we will implement a post-double-selection LASSO procedure to identify significant baseline predictors of the outcome and treatment. The final control set will consist of the union of theoretically motivated covariates and those selected by LASSO.

When using z-scores we will use OLS. Alternatively, we will also use the raw scores and employ fractional regression to check for robustness.

7.2 Difference-in-Differences Specification

We will also estimate treatment effects using a difference-in-differences specification in first differences:

$$\Delta Y_{iv} = Y_{iv,1} - Y_{iv,0} = \alpha + \beta\, Treatment_v + \gamma' X_{iv,0} + \eta_{iv}.$$

This specification will estimate treatment effects using changes between baseline and endline. Under random assignment, both the ANCOVA and difference-in-differences estimators are consistent.

Standard errors in this specification will also be clustered at the village level.

Treatment will be constructed based on the pre-specified contrasts:

1. Control vs Any Treatment
2. Control vs T1
3. Control vs T2
4. T1 vs T2

7.3 Sensitivity Analysis

We will conduct the following pre-specified sensitivity analyses.

7.3.1 Excluding Control Women Who Own Smartphones at the Endline

We will redefine the control group to exclude women who report owning a smartphone between the baseline and the endline. The restricted control group will be compared to each treatment arm and to the pooled treatment group.

Specification remains:

$$Y_{iv,1} = \alpha + \beta Treatment_v + \rho Y_{iv,0} + \gamma' X_{iv,0} + \varepsilon_{iv}$$

but estimated on the restricted sample.

7.3.2 Excluding Control Women Who Own Any Phone

We will further restrict the control group to women who report owning neither a smartphone nor a keypad phone at the endline. This specification tests whether treatment effects are driven by general phone access rather than smartphone-specific exposure.
The same ANCOVA specification will be estimated on this restricted sample.

7.3.3. Keypad Phone Sensitivity

To examine whether baseline keypad phone ownership drives results, we will:

1. Remove control-group women who acquired keypad phones after baseline.
2. Construct an indicator variable equal to one if a control-group woman owned a keypad phone at baseline and zero otherwise.
3. Include this baseline keypad indicator in the ANCOVA specification:

$$Y_{iv,1} = \alpha + \beta Treatment_v + \rho Y_{iv,0} + \delta Keypad_{iv,0} + \gamma' X_{iv,0} + \varepsilon_{iv}$$

We will test whether δ is statistically significant and whether inclusion of this indicator meaningfully alters the estimated treatment effects.

7.4 Instrumental Variables

Actual Smartphone ownership is a post-treatment outcome and depends on adherence to treatment assignment. This adherence may not be random. We define smartphone ownership as an indicator equal to one if the woman owns a smartphone at endline and zero otherwise. Because ownership may be influenced by unobserved characteristics correlated with empowerment outcomes, directly including this variable in the primary ITT specification would raise endogeneity concerns.
We therefore estimate secondary instrumental variables (IV) specifications in which randomized treatment assignment serves as an instrument for smartphone ownership. The first-stage equation is:

$$Smartphone\ ownership_{iv} = \pi\, Treatment_v + \theta' X_{iv,0} + u_{iv},$$

and the second-stage equation is:

$$Y_{iv,1} = \alpha + \beta\, \widehat{Smartphone\ ownership}_{iv} + \gamma' X_{iv,0} + \varepsilon_{iv}.$$

Under the standard IV assumptions, including relevance and the exclusion restriction that treatment assignment affects outcomes only through smartphone ownership, these estimates identify the Local Average Treatment Effect (LATE) for compliers.
IV specifications will also be estimated within baseline phone-constrained subgroups to examine heterogeneity in compliance.

---

8. Multiple Hypothesis Testing

We will report for the various secondary outcomes:

- Romano–Wolf stepdown adjusted p-values (FWER control), and
- Benjamini–Hochberg q-values (FDR control).

The composite empowerment index is the primary outcome and excluded from MHT adjustment.
MHT will be applied within each pre-specified contrast family.

---

9. Attrition

We will:

1. Test for differential attrition across treatment arms.
2. Find out if the attrition population differ on baseline characteristics.
3. If attrition differs, compute Lee (2009) bounds for ITT estimates under monotone selection.

Lee bounds will be applied to ITT and ATT estimates.

---

10. Secondary Outcomes

The following outcomes are secondary and will be analyzed using the ITT specification in Section 7. These are conceptually distinct from the primary empowerment index. These variables will also be standardized relative to the control group.

10.1 Financial Well-being Index

Items will combine three items reflecting whether the respondent feels that her financial situation prevents her from having the things she wants in life, whether she is just getting by financially, and whether she is concerned that her current or future savings will not last. Responses will be coded so that higher values indicate greater financial wellbeing (i.e., lower perceived financial stress), standardized, and averaged to form the index in [0,1].

10.2 Intra-Household Bargaining Power (IHBP)

We will use willingness-to-pay (WTP) as a proxy for bargaining power.

- WTP ranges from 0 to 300.
- Higher WTP indicates lower bargaining power.
- We will construct:

$$IHBP_i = 1 - \frac{WTP_i}{WTP\ max_i}$$

so that higher values correspond to higher bargaining power and the measure lies in [0,1].

10.3 Mental Health Index

Index will combine two measures reflecting core symptoms of psychological distress: reduced interest or pleasure in activities that were previously enjoyable, and feelings of sadness, depression, or hopelessness over the past two weeks. Responses will be coded so that higher values indicate worse mental health in [0,1].

10.4 Multidimensional Poverty Index (MPI)

We will construct a Multidimensional Poverty Index (MPI) following the Alkire–Foster (2011) methodology. The index will incorporate deprivations across multiple dimensions of child wellbeing, housing quality, asset ownership, and hygiene practices. Specifically, the education dimension will capture whether all children aged 4–6 attend pre-school and whether all children aged 7–14 attend school. The housing dimension will be based on the enumerator-assessed condition of the dwelling (pucca, semi-pucca, or kutcha). The assets dimension will include ownership of durable goods such as television, bicycle, motorbike, car, tractor or truck, computer, refrigerator, and washing machine. The hygiene dimension will reflect the frequency with which children wash their hands with soap before meals and after defecation.

Each indicator will be converted into a binary deprivation measure using pre-specified cutoffs. Dimension-specific weights will be assigned, and a household will be classified as multidimensionally poor if its weighted deprivation score exceeds a pre-defined poverty threshold 0.3. This will serve as an objective measure of deprivation in [0,1], where high values indicate high poverty.

10.5 Smartphone Norms Outcomes
(A) Recharge Behavior
Binary variable:

- 1 if respondent recharged internet herself in the past month before survey
- 0 otherwise

A similar variable will be created based on what they reported they would do regarding recharge, when asked during the monthly survey.
We will estimate the effect of the smartphone social norms correction digital intervention (cross randomized within T1 and T2 villages) on these outcomes.
(B) Norms Beliefs
We will measure attitudes and perceived norms surrounding women's smartphone ownership using five related statements. These capture the respondent's own belief that women should have smartphones, her perception of her husband's view, the perceived views of members of her in-laws' household, and broader community norms, including what most people and most women in the village are believed to think. Together, these items reflect both personal support and perceived social approval for women's access to smartphones. Responses will be coded so that higher values indicate stronger normative support and averaged to form the index in [0,1]. We will test if this norms index is shifted by the smartphone social norms correction digital intervention.
(C) Demand for Private money
This variable is based on a willingness to pay measure, where the women decide between keeping a smaller amount for themselves or allocating a larger amount to their husband. We will test if demand for private money changes due to the smartphone social norms correction digital intervention.
(D) Check for interaction
We will also plan to check if there is any interaction effect arising from the norms correction and curated content on the outcomes of interest.

10.6 Subjective Deprivation
We will include a subjective measure of perceived living standards using a self-anchoring ladder scale. Respondents will be asked to place themselves on a ladder representing the distribution of living standards, where one end corresponds to the worst-off individuals and the other to the best-off. This measure captures perceived relative economic position.
In addition, respondents will be asked how many steps they would like to move up the ladder over the next ten years, providing a measure of economic aspirations. These variables will be used to assess whether subjective perceptions of wellbeing align with the multidimensional poverty index and whether treatment exposure shifts aspirations over time.

10.7 Consumption
We will combine self-reported household consumption expenditures across a range of goods and services. These data will be aggregated to construct a measure of total household consumption, which will be used as a secondary outcome to in the ITT specification.

10.8 Social Comparison Index
These include whether respondents compare their financial situation or possessions to women they see online whether they feel worse about themselves after observing women with higher education or English skills, whether exposure to successful women on social media inspires them to improve their lives, and whether social media creates pressure to spend money or change how they dress. Items will be reverse coded where necessary so that higher values consistently reflect more negative responses. The components will then be averaged to construct an index scaled to lie in [0,1].

10.9 Body Esteem Index
These include whether the respondent feels pressured to look a certain way because of women seen on social media, whether she likes what she sees when she looks in the mirror, and whether she feels satisfied with her body. The question on appearance pressure will be asked only of respondents who report using social media. Items will be reverse coded where appropriate so that higher values consistently reflect more positive body image and lower perceived pressure. Responses will be normalized and averaged to construct an index scaled to lie between 0 and 1.

---

11. Mechanism Analysis (Exploratory)
Mechanism analyses are pre-specified as exploratory and will be interpreted as descriptive correlations.

11.1 Engagement at endline
We will examine engagement with intervention content as a potential mechanism linking treatment assignment to outcomes. For respondents in treatment arms that received curated WhatsApp content, we will measure how frequently they report watching the YouTube videos shared through the village WhatsApp group and how useful they perceived those videos to be. In addition, for treatment arms that received surveys via WhatsApp, we will record who completed the surveys on the respondent's phone. These measures capture intensity of exposure and degree of direct engagement with the intervention, allowing us to assess whether treatment effects operate through content consumption and active participation.

11.2 Enablers
We will measure skill acquisition as a potential enabling mechanism by asking respondents what new skills they have learned from watching videos on a smartphone. The question will cover a range of domains, including cooking new recipes, sewing and tailoring, teaching children, beauty and parlor techniques, online financial transactions, handicrafts, and creating and uploading online content. Responses will capture both practical livelihood skills and digital or financial capabilities. We will construct an indicator for having learned any new skill, and, where relevant, summary measures reflecting the number or type of skills acquired. These measures will be used to assess whether smartphone exposure translated into tangible capability gains that may support empowerment outcomes.

11.3 Barriers
We will measure barriers to smartphone ownership and use among women in the treatment arms to understand constraints that may limit engagement. Respondents will be asked to report the challenges they faced in owning or using a smartphone, including social taboos, disapproval from

husbands or in-laws, children taking the phone, literacy difficulties, language barriers, lack of time, concerns about fraud or financial loss, or perceptions that the phone is not useful. These responses will capture social, household, and capability-related barriers.

In addition, we will measure the level of family support for smartphone use by asking respondents to assess the extent to which their husband, in-laws, and other family members support their use of the device. Together, these measures will provide insight into enabling and constraining household dynamics that may shape treatment engagement and downstream empowerment outcomes.

11.4 Time Use

We will measure smartphone usage intensity, using the endline as well as monthly data, as a potential mechanism linking treatment assignment to outcomes. Respondents will report the amount of time they spend on social media applications such as Facebook, Instagram, or YouTube on a typical day, as well as their average daily smartphone usage. These measures capture both content-specific exposure (social media use) and overall digital engagement (total smartphone use). Responses will be coded into ordered categories and, where appropriate, normalized to construct summary measures of usage intensity. These variables will be used to assess whether treatment effects operate through increased time spent using smartphones and digital platforms.

11.5 Smartphone Exposure and Usage Index

We will construct a Smartphone Exposure Index to capture digital access, autonomy, literacy, and intensity of use. The index will combine measures of shared or personal phone access, ability to search for information online, frequency and autonomy of smartphone use, use of key applications (including communication, social media, financial, and informational apps), engagement with government schemes, and awareness of digital security features. Items will be coded and normalized so that higher values reflect greater exposure.

11.6 Employment and Employability Index

We will measure labor market engagement as a potential mechanism through which smartphone access may influence economic empowerment. These measures will capture whether the respondent's income has increased in the past nine months, whether she is currently engaged in paid work, and, if not employed, whether she is actively seeking or available for work. For those not currently working, we will also record efforts undertaken to secure employment, including registration with government or private employment agencies, engagement with panchayat offices, online job search, or other job-seeking activities. Together, these variables will allow us to assess whether treatment exposure translates into greater labor force participation, income growth, or increased job search intensity. We will additionally examine whether online job search differs between T1 and T2.

11.7 Human Capital Enhancement Index

We will measure digital capability and institutional engagement as mechanisms reflecting human capital enhancement. These measures capture whether smartphone access translates into improved digital literacy, functional use of applications, financial capability (e.g., UPI and mobile banking), skill acquisition, job search, and engagement with government schemes. Together, they reflect investments in knowledge, skills, and productive capacities that may mediate the impact of treatment on empowerment outcomes.

11.5 Qualitative Insights

We collect insights from in-depth interviews to explore the mechanisms through which smartphone ownership may influence women's empowerment and to identify barriers that may limit its impact. The discussions will examine whether ownership itself drives empowerment or whether effects operate through intermediate channels such as improved access to information, financial inclusion, job opportunities, enhanced confidence, stronger bargaining power, and expanded social networks. We will also probe key barriers, including socio-cultural restrictions, lack of digital skills or confidence, and time constraints.

Separate discussions will be held with women from Treatment Group 1, Treatment Group 2, and the control group. For treatment groups, we will explore patterns of smartphone use, digital literacy, perceived benefits, changes in household decision-making roles, engagement with intervention content (where applicable), and experiences of restrictions or support from family members. For control group women, we will focus on access constraints, norms around women's smartphone ownership, and anticipated benefits or challenges.

These qualitative insights will complement the quantitative analysis by helping interpret observed effects, testing hypothesis where feasible and identifying plausible mechanisms underlying empowerment outcomes.

11.6 Spillovers Within Household

We will examine whether smartphone ownership generates intra-household spillover effects. Specifically, respondents will be asked whether any household members (e.g., husband, children, in-laws) benefited after the respondent owned and used a smartphone. These responses will be used to construct spillover indicators and summary measures.

To assess whether the video-based treatment arm generates greater spillover effects relative to basic smartphone provision, we will estimate the following ITT specification:

$$Spillover_{iv} = \alpha + \beta_1 T1_v + \beta_2 T2_v + \gamma' X_{iv,0} + \varepsilon_{iv}$$

and formally test whether $\beta_1 = \beta_2$. Standard errors will be clustered at the village level.

In addition, we will examine whether spillover measures are associated with empowerment outcomes and smartphone exposure to assess whether intra-household benefits constitute a potential mechanism. These analyses will be interpreted as exploratory and mechanism-consistent rather than definitive causal mediation tests.

11.7 Time use from monthly surveys

Using high-frequency monthly survey data, we will construct time-use measures capturing the amount of time respondents spend searching for or accessing information on their smartphones across specific domains. These domains include: (i) general and digital information search, (ii) communication and entertainment, (iii) human capital development and income-generating activities, (iv) health and hygiene-related information, (v) child education support, and (vi) awareness of and search for government schemes.

For each domain, we will compute time-based measures reflecting intensity of search and engagement. These measures will be used to assess whether treatment effects operate through increased time allocation toward productive, informational, and capability-enhancing activities on the smartphone.

13. Monthly Panel Analysis

Using the high-frequency monthly survey data collected for Treatment Group 1 (T1) and Treatment Group 2 (T2), we will exploit the panel structure to study the dynamics of smartphone use and related mechanism outcomes. These data allow us to track, at the individual level, changes in time spent on different activities, the composition of apps used, and confidence in smartphone use over time.

Dynamic treatment contrast (T2 vs T1 over months)

First, we will estimate dynamic treatment contrasts to assess whether outcomes in the video-based arm (T2) diverge from T1 over time. Specifically, we will estimate specifications with month fixed effects and treatment-by-month interactions, allowing us to trace the evolution of differences in usage intensity, human capital–related search, financial transactions, and digital confidence. These models will identify whether repeated exposure to curated video content leads to gradual or cumulative behavioral change.

$$Y_{itm} = \alpha + \sum_{m=1}^{M} \beta_m \ (T2_i \times \mathbf{1}\{t = m\}) + \lambda_m + \gamma' X_{i0} + \varepsilon_{itm},$$

where Y_{itm} is a monthly outcome (time-use domain, app-composition, confidence), λ_m are month fixed effects, and X_{i0} are baseline controls. Inference clustered at village.

Cumulative exposure

Second, we will estimate cumulative exposure (dose–response) models in which monthly outcomes are regressed on the number of months of video exposure received up to period t. This approach allows us to test whether behavioral changes scale with sustained exposure to digital content.

Let $CumExp_{im}$ be cumulative months of exposure to video content up to month m (for T2; 0 for T1), or cumulative number of videos sent up to month m.

$$Y_{im} = \alpha + \beta \, CumExp_{im} + \lambda_m + \gamma' X_{i0} + \varepsilon_{im}.$$

Persistence

Third, we will examine persistence and habit formation by studying whether increased usage in earlier months predicts continued engagement in later months, and whether T2 shifts the trajectory of usage or confidence relative to T1.

$$Y_{im} = \alpha + \rho \, Y_{i,m-1} + \beta \, T2_i + \lambda_m + \gamma' X_{i0} + \varepsilon_{im}.$$

Individual fixed effects (within-person dynamics)

Fourth, where appropriate, we will use individual fixed-effects models to exploit within-person variation over time, thereby controlling for time-invariant heterogeneity in baseline ability, motivation, or household characteristics.

$$Y_{im} = \alpha_i + \lambda_m + \beta \, T2_i + \varepsilon_{im},$$

Finally, we will explore descriptive correlates of usage and confidence growth, including whether the respondent is the primary phone user, reported difficulties in using the phone, and internet connectivity constraints.

---

15. Heterogeneity Analysis

We pre-specify heterogeneity along baseline characteristics including:

- Percentage of sons
- Trust
- English Proficiency
- Education of own and husband

- Caste
- Low vs high usage (more than 1 hours spent on phone on a typical day)
- Low vs high engagement (replied to 9 or more monthly surveys)
- Baseline Phones owned or not
- Telephonic vs physical survey

We will also estimate heterogeneous treatment effects using causal forest methods. The causal forest will be used to identify predictors of treatment effect heterogeneity; results are exploratory and will not be used to redefine primary estimands.

---

16. Inference Procedures

We will report conventional cluster-robust standard errors clustered at the village level. In addition, given the cluster-randomized design and the possibility of a limited number of clusters, we will complement asymptotic inference with two design-consistent and small-sample–robust procedures.

Randomization Inference (RI). We will conduct randomization inference by repeatedly reassigning treatment at the village level following the original randomization protocol (preserving the number of treated villages). For each re-randomization, we will re-estimate the pre-specified ITT coefficient for the corresponding contrast and construct an empirical reference distribution of the test statistic. RI p-values will be computed as the share of re-randomizations producing a statistic at least as extreme as the observed estimate (two-sided).

Wild Cluster Bootstrap. We will also report p-values based on the wild cluster bootstrap procedure, which is designed to improve finite-sample inference with clustered data.