Supplement 1: Sampling strategy and sample size considerations for the study

Given that the design and implementation strategy of Project Samuday precluded a cluster-randomized trial design, the next best alternative was determined to be a “difference-in-difference” (DID) design with matching at the gram panchayat level between intervention areas in Hardoi and control areas in the adjacent district of Sitapur. Since the Samuday project intervention is focused on three blocks in Hardoi district, gram panchayats (GPs) from three socio-demographically similar census blocks in the adjacent Sitapur district were selected to serve as controls, separated by a “buffer layer” of census blocks to reduce the possibility of geographic spillover effects.

The intervention census blocks in Hardoi contain an average of 58 gram panchayats (GP), each with a total of 173 GPs in the three blocks (41 in Kachhauna; 71 in Behadar; 61 in Kothwan). The control census block in Sitapur contains an average of 83 GPs, each with a total of 250 GPs in the three blocks. Each GP typically includes one to three villages, which average 300-350 households per village. Each of the 173 GPs in the intervention arm was pair-matched with a comparison GP in the control blocks using a method called “coarsened exact matching” (CEM) in order to maximize the balance of a set of key sociodemographic covariates.

We calculated the sample size required to detect the difference in the difference between intervention and control arm for each of a series of key indicators and achieve 80% power with a type 1 error of 0.05. The sample size incorporates a design effect ranging from 1.7 – 2.2 and a non-response rate of 10 percent. The table below presents sample size scenarios for these indicators, assuming a 6% - 11% difference in the difference between intervention and control arms. This assumes a 10% and 15% difference in indicators for the intervention arm between baseline and end-line, and a 4% difference in indicators for the control arm between the baseline and end line. Thus, for intervention and control areas, a total sample size of 12854 households (3146 in each study arm for each round) was determined to be sufficient to detect a 6% difference in difference in all indicators except for institution delivery in the 24 months before the survey, for which an 11% difference in difference can be detected. All head of households and women 15-49 years old were interviewed, and anthropometric data (weight and height) was collected from children under-five. Based on a crude annual birth rate of 28.1 per 1000 population and a mean household size of 5.5, it was expected that approximately 15016 women aged 15-49 would be interviewed, and 13096 children under the age of five would be measured.

Before conducting the survey, a mapping and listing exercise was conducted in the selected primary sampling units (PSU). During the mapping and listing exercise, the “Right-Hand Rule” was taken after starting from any important landmark of the PSU. From the list of all households, we randomly selected 25 households as the potential list of participants, from where the first 18 households were selected for interview skipping any house where no person is living. In any case, if the household head was absent, the interviewer attempted to complete the interview by returning twice in the same household in the next two days. This comprehensive procedure helped the survey to achieve an overall response rate of over 99%.
Table: 1: sample size considerations for the study

| Type of respondent | Indicator | Intervention arm (Baseline prevalence) | Intervention arm (Endline prevalence) | Endline – Baseline difference in control arm | Difference in Difference | Sample size per arm per round (Design effect) | Baseline sample size | Sample size per cluster/GPS |
|-------------------|-----------|----------------------------------------|----------------------------------------|-------------------------------------------|--------------------------|---------------------------------------------|---------------------|----------------------------|
| Heads of household | Household with improved sanitation facility | 23 | 33 | 4 | 6 | 3146 (1.7) | 6292 | 18 |
| | Household with any usual member covered by a health scheme or health insurance | 5 | 15 | 4 | 6 | 1239 (1.7) | 2478 | 7 |
| Reproductive age women | With 10 or more years of schooling | 20 | 30 | 4 | 6 | 3754 (2.2) | 7508 | 22 |
| | Participation in women’s group | 15 | 25 | 4 | 6 | 3142 (2.2) | 6284 | 18 |
| | Institutional delivery | 60 | 75 | 4 | 11 | 1408 (2.2) | 2816 | 8 |
| Children <5y | Percent wasted (weight for height) | 16 | 26 | 4 | 6 | 3274 (2.2) | 6548 | 19 |
Supplement 2: Power calculation for assessing the adequacy of the sample

During the baseline evaluation of Project Samuday, 6,218 households participated in the survey with a response rate of >99%. Power analysis is conducted to determine if the sample size of the study is adequate for statistical analyses and for appropriate generalization to the population. While the sample size of this study was estimated to detect changes in several indicators mentioned in the Supplement 1, it is necessary to estimate if we have sufficient power to statistically measure the proportion of the tobacco and generalize the findings.

The following table provides the result of the power analysis, and the estimates were used for the analysis. The result suggested that for all three types of samples – total, men, and women – the analysis will have sufficient power (> 80%) for an estimate of the average proportion of tobacco users in Uttar Pradesh.

Table 2: Power calculation for assessing the adequacy of the sample

| Type of sample | Significance (alpha) | Number of Clusters (K) | Average number of respondents (M) | Total Sample Size (N = K x M) | GATS2 reported proportion (m0) | Study reported proportion (ma) | Difference in prevalence (delta = ma – m0) | Study reported intraclass correlation (rho) | Power  |
|---------------|----------------------|------------------------|----------------------------------|-------------------------------|-----------------------------|---------------------------------|------------------------------------------|------------------------------------------|--------|
| Total         | 0.05                 | 346                    | 18                               | 6,228                         | 0.355                       | 0.6246                          | 0.2696                                   | 0.03503                                 | 1.00   |
| Male          | 0.05                 | 346                    | 15.4                             | 5,328                         | 0.521                       | 0.7065                          | 0.1855                                   | 0.03931                                 | 1.00   |
| Female        | 0.05                 | 309                    | 3                                | 927                           | 0.177                       | 0.1446                          | -0.03241                                 | 0.8011                                   |        |

Note: The estimates for GATS 2 in Uttar Pradesh: https://tmc.gov.in/images/act/Uttar%20Pradesh%20GATS-2%20Factsheet%20.pdf

Table 3: Stata Code for the Power Calculation:

| Type of sample | Stata Code                                                                 |
|---------------|---------------------------------------------------------------------------|
| Total sample  | power oneproportion 0.355 0.6246381, k(346) m(18.0) rho(.0350276) table |
| Male          | power oneproportion 0.521 0.7065136, k(346) m(15.4) rho(.0393123) table   |
| Female        | power oneproportion 0.177 0.1445916, k(309) m(3) rho(0) table            |
## Supplement 3: Table 4 - Description of the explanatory variables

| Variables                      | Description                                           | Type                                      |
|-------------------------------|-------------------------------------------------------|-------------------------------------------|
| **Individual demography**     |                                                       |                                           |
| Age                           | Self-reported age of the respondents                  | Five: Categories                          |
|                               |                                                       | - ≤ 30 years                              |
|                               |                                                       | - 31-40 years                             |
|                               |                                                       | - 41-50 years                             |
|                               |                                                       | - 51-60 years                             |
|                               |                                                       | - > 60 years                              |
| Gender                        | Self-reported age of the respondents                  | Two Categories                            |
|                               |                                                       | - Male                                    |
|                               |                                                       | - Female                                  |
| Religion                      | Self-reported religion of the household               | Two Categories                            |
|                               |                                                       | - Hindu                                   |
|                               |                                                       | - Muslim and others                       |
| Caste                         | Self-reported social caste of the household           | Three Categories                          |
|                               |                                                       | - General                                 |
|                               |                                                       | - ST/SC                                   |
|                               |                                                       | - OBC and others                          |
| Marital Status                | Self-reported marital status of the respondent        | Three Categories                          |
|                               |                                                       | - Never married/Not stated                |
|                               |                                                       | - Married                                 |
|                               |                                                       | - Widow/Divorced/Separated                |
| Education                     | Self-reported educational attainment of the respondent| Four Categories                           |
|                               |                                                       | - Illiterate                              |
|                               |                                                       | - Up to primary (5th grade)               |
|                               |                                                       | - Up to secondary (10th grade)           |
|                               |                                                       | - Above secondary                         |
| Occupation                    | Self-reported occupation of the respondent            | Six Categories                            |
|                               |                                                       | - Cultivator                              |
|                               |                                                       | - Wage laborer                            |
|                               |                                                       | - Self-employed & Others                  |
|                               |                                                       | - Salaried worker                         |
|                               |                                                       | - Housewife                               |
|                               |                                                       | - Unemployed                              |
| Household size                | Self-reported number of a household member living in  | Continuous with three spline terms:      |
|                               | the house for the last six months                    | - Up to 3 members                         |
|                               |                                                       | - Up to 10 members                        |
|                               |                                                       | - More than ten members                   |
| Variables | Description | Type |
|-----------|-------------|------|
| **Household wealth (assets quintile)** | Asset index developed by principal component analysis using 27 binary variables. The variables include information regarding ownership of household assets, house, and land ownership. The standardized score of the first component was used to create five asset quintile groups, where Quintile 1 was assigned to the least wealthy household, and Quintile 5 was assigned to the most wealthy household. | Five Categories: From Quintile 1 to 5 |
| **Individual personality traits** | | |
| Freedom of making decisions | Derived from the question: “How much freedom do you have in making personal decisions?” Two separate categories were developed from the original responses: 1. Low = “No freedom at all,” “Freedom in very few decisions,” and “Freedom in some decisions.” 2. High = “Freedom in most decisions” and “Freedom in all decisions.” | Two Categories |
| Satisfaction with life circumstances | Tertile developed from the first component of the principal component analysis using 19 binary variables related to household head’s satisfaction towards minimum needs in the following areas: daily food, meals in holidays, clothing, shoes, accommodation, water, electricity, furniture, personal hygiene products, transportation, education etc. | Three Categories |
| Level of happiness | Derived from the question: “Taking all things together, would you say you are happy, unhappy or neither?” Three separate categories were developed from the original responses: 1. Unhappy = “Very unhappy” and “Somewhat unhappy” 2. Neither happy nor unhappy = Neither happy nor unhappy 3. Happy = “Somewhat happy” and “Very happy.” | Three Categories |
| **Perceived accessibility** | Perceived accessibility was measured by household head’s self-reported perception of improvement of village infrastructure service: “How has the functioning of infrastructure in your village (e.g., roads, electricity, and water supply) changed since last year?” | 3 Categories |
| **Social Capital** | | |
| Individual Level Social Capital | Twelve modified items of Adapted Social Capital Assessment Tool-India (SASCAT-I)² reported by 6,218 household heads were used to perform a multilevel Confirmatory factor analysis (MCFA). Four uniquely identified factor emerged from the MCFA model at both individual and the community (PSU level). | |
| Community engagement | Standardized factor score derived from the MCFA model with three indicators: 1. Group Membership = In the last 12 months, participated in or received any benefit from any community group 2. Collective Action = In the last 12 months, worked together with other community members and attempted to address a problem or common issue of the village 3. Development Discussion = In the past 12 months, spoke with anyone about the development of the village | Binary |
| Social Support | Standardized factor score derived from the MCFA model with three indicators: 4. Emotional Support = In the last 12 months, received any emotional, social support 5. Financial Support = In the last 12 months, received any financial social support 6. Informational Support = In the last 12 months, received any informational social support | Binary |
| Trust | Standardized factor score derived from the MCFA model with two indicators: 7. Trust in Leaders = Overall, trust in village leaders 8. Trust in Strangers = Overall, trust in unfamiliar people residing in the village 9. Trust in Neighbors = Overall, trust in village neighbors | 3 Categories |
| Variables                  | Description                                                                 | Type                              |
|---------------------------|-----------------------------------------------------------------------------|-----------------------------------|
| Social cohesion           | Standardized factor score derived from the MCFA model with four indicators: | 3 Categories                      |
|                           | 10. Social Harmony = People in this village generally have good relationships with each other | - No                              |
|                           | 11. Sense of Belonging = Feel that you belong to this village                | - Sometimes                       |
|                           | 12. Sense of Fairness = People in this village would try to take advantage of you if they get the chance | - Yes                             |
| Community Level Social Capital | PSU level standardized factor score derived from level 2 of Multilevel MCFA | Continuous                        |
| Community engagement      | PSU level standardized factor score derived from level 2 of Multilevel MCFA | Continuous                        |
| Social support            | PSU level standardized factor score derived from level 2 of Multilevel MCFA | Continuous                        |
| Trust                     | PSU level standardized factor score derived from level 2 of Multilevel MCFA | Continuous                        |
| Social cohesion           | PSU level standardized factor score derived from level 2 of Multilevel MCFA | Continuous                        |
| Social Environment        | PSU level standardized factor score derived from level 2 of Multilevel MCFA | Continuous                        |
| Community demography      | Gram Panchayat Size = Tertile developed based on the population of the gram panchayat reported from 2011 Census of India | 3 Categories                      |
|                           | Community wealth = Average scores of the first component of principal component analysis from the households of each cluster. The score is standardized for easier interpretation | Continuous                        |
| Community health service function | Average cluster score of individual’s perceptions on the improvement of community health services | Continuous                        |
| Community infrastructure improvement | Average cluster score of individual’s perceptions on the improvement of village infrastructure service | Continuous                        |
| Community tobacco consumption | Scaled no-self cluster proportion of tobacco use was generated by calculating the proportion of the household heads in the community (PSU) who consumed tobacco while excluding the respondent both from the numerator and denominator and then multiplying the proportion by 10. One unit increase in of this scaled indicator represents a 10% increase in “Non-self” cluster proportion of Tobacco use | Continuous                        |

Note: ST/SC = Scheduled castes and scheduled tribes, OBC = Other backward castes, MCFA= Multilevel confirmatory factor analysis
1 = Hasan, M.Z., Leoutsakos, J.-M., Story, W., Dean, L.T., Rao, K.D., Gupta, S., 2019. Exploration of Factor Structure and Measurement Invariance by Gender for a Modified Shortened Adapted Social Capital Assessment Tool in India. Frontiers in Psychology 10, 2641.
2 = De Silva MJ, Harpham T, Tuan T, Bartolini R, Penny ME, Huttly SR. Psychometric, and cognitive validation of a social capital measurement tool in Peru and Vietnam. Social Science & Medicine. 2006 Feb;62(4):941–53.
3 = Office of the Registrar General & Census Commissioner. States Census 2011 [Internet]. 2011 [cited 2017 Dec 13]. Available from: http://www.census2011.co.in/states.php
Supplement 4: Peer influence of tobacco use measured by non-self cluster proportion of tobacco use

To account for the endogeneity, we created a “non-self cluster proportion of tobacco use” instead of measuring the absolute proportion of tobacco users in the community. Generally, the absolute proportion of tobacco user in the community is measured by:

\[
\text{Proportion of tobacco user in a cluster} = \frac{\text{Number of tobacco users in a cluster}}{\text{Total number of respondents in the cluster}}
\]

However, for measuring the “non-self cluster proportion,” we have to calculate separate scores for each household head based on their smoking status.

If a household head is a tobacco user, then:

\[
\text{Non-self score of tobacco use for tobacco user} = \frac{\text{Number of tobacco users in a cluster} - 1}{\text{Total number of respondents in the cluster} - 1}
\]

Furthermore, if a household head is not a tobacco user, then:

\[
\text{Non-self score of not tobacco use for tobacco user} = \frac{\text{Number of tobacco users in a cluster}}{\text{Total number of respondents in the cluster} - 1}
\]

Lastly, the score of all household heads (users and non-users) were averaged at the community or PSU level. This alternative calculation of cluster proportion of tobacco use accounts for the contribution of each respondent by excluding their contribution while measuring the proportion.
Supplement 5: Confirmatory factor analysis to generate social capital measures for the household head

Adapted from original SASCAT (De Silva et al., 2006) and SASCAT-B (Story et al., 2015) using rapid cognitive interviewing, the SASCAT-I is a self-reported measure of social (Hasan et al., 2019). The study sample household heads from 6,218 randomly selected households who responded to the SASCAT-I during a community-level multistage cross-sectional survey. To assess the factor structure of social capital, 12 binary items were generated from 13 self-reported questions of SASCAT-I. The first seven questions of the tool are related to structural social capital - Group membership (2 questions), Collective action (2 questions), Social support (3 questions). The last six questions are related to cognitive social capital- Trust (3 questions) and Social Cohesion (3 questions). Four unique factors were identified both at the level of individual household head and at the community (PSU) level during Horn’s (1965) parallel analysis.

Figure 1: Scree plots indicating the possible number of factors identified at individual and community level for household heads using SASCAT-I

Next, the factor structure of household head’s social capital was identified through multilevel exploratory factor analysis using “Weighted Least Square Mean and Variance” (WLSMV) adjusted estimator using a polychoric correlation matrix and holding factor variances fixed to one (Heck and Thomas, 2015). To assess the construct validity of the social capital factor structure and to generate standardized factor scores of each construct of social capital, we
conducted a multilevel confirmatory factor analysis (MCFA), which used “Maximum Likelihood Robust” estimator and polychoric correlation matrix, holding factor variances fixed to one. The model presented an adequate fit with the data with RMSEA = 0.32, CFI = 0.91, TLI = 0.88, SRMR = 0.05, $\chi^2$ value = 715.32, df = 96, p < 0.01. Figure 2 presents the path diagram of MCFA with unstandardized factor loadings and inter-factor correlations. Associated factor scores of the eight social capital constructs were extracted from the MCFA and used in the regression models to assess the relationship between each construct of social capital and tobacco use.

**Figure 2: Path diagrams presenting unstandardized factor loadings and inter-factor correlations of four-factor multilevel CFA model for household heads in rural Uttar Pradesh (n = 6,218)**

*Note: Factors: OP = Organizational participation, SS = Social Support, TR = Trust, SC = Social Cohesion
Goodness of fit Indices (of the same model estimated by WLSMV): RMSEA = 0.32, CFI = 0.91, TLI = 0.88, SRMR = 0.05, $\chi^2$ value = 715.32, df = 96, p < 0.01
GM = Group Membership, CA = Collective Action, DD = Development Discussion, ES = Emotional Support, FS = Financial Support, IS = Informational Support, TL = Trust in Leaders, TS = Trust in Strangers, TN = Trust in Neighbors, SH = Social Harmony, SB = Sense of Belonging, SF = Sense of Fairness*
References:
Hasan, M.Z., Leoutsakos, J.-M., Story, W., Dean, L.T., Rao, K.D., Gupta, S., 2019. Exploration of Factor Structure and Measurement Invariance by Gender for a Modified Shortened Adapted Social Capital Assessment Tool in India. Front. Psychol. 10, 2641.

Heck, R.H., Thomas, S.L., 2015. An introduction to multilevel modeling techniques: MLM and SEM approaches using Mplus. Routledge.

Horn, J.L., 1965. A rationale and test for the number of factors in factor analysis. Psychometrika 30, 179–185. https://doi.org/10.1007/BF02289447

Silva, M.J.D., Harpham, T., Tuan, T., Bartolini, R., Penny, M.E., Huttly, S.R., 2006. Psychometric and cognitive validation of a social capital measurement tool in Peru and Vietnam. Soc. Sci. Med. 62, 941–953. https://doi.org/10.1016/j.socscimed.2005.06.050

Story, W.T., Taleb, F., Ahasan, S.M., Ali, N.A., 2015. Validating the Measurement of Social Capital in Bangladesh A Cognitive Approach. Qual. Health Res. 25, 806–819.
Supplement 6: Figure 3 – Distribution of individual and community social capital standardized factor score among household heads based on tobacco consumption status in rural Uttar Pradesh, India (n = 6,218)
Supplement 7: Bivariate distribution of individual social capital standardized factor score among men and women household heads in rural Uttar Pradesh, India

Table 5: Bivariate distribution of individual social capital standardized factor score among men and women

| Individual Social Capital standardized factor score | Gender of the respondent | Total (n = 6,218) |
|---------------------------------------------------|--------------------------|-----------------|
|                                                   | Men (n = 5,312)          | Women (n = 906) | P values |
| Community engagement                              | 0.01                     | -0.08           | 0.01     | 0.00 |
| Social support                                    | 0.02                     | -0.09           | 0.00     | 0.00 |
| Trust                                             | -0.01                    | 0.07            | 0.02     | 0.00 |
| Social cohesion                                   | -0.01                    | 0.04            | 0.21     | 0.00 |

| Gender of the respondents who consumed tobacco products |
|----------------------------------------------------------|
| Individual Social Capital standardized factor score | Men (n = 3,753) | Women (n = 131) | P values |
|                                                   | Mean                     | Mean             | Mean    |
| Community engagement                              | 0.01                     | 0.11             | 0.28    | 0.01 |
| Social support                                    | 0.01                     | 0.11             | 0.25    | 0.01 |
| Trust                                             | -0.01                    | 0.10             | 0.21    | -0.01 |
| Social cohesion                                   | -0.01                    | 0.05             | 0.52    | -0.01 |
Supplement 8: Figure 4 – **Pearson correlation matrix representing the correlation between community-level social capital measures and peer (normative) influence for men (lower triangle) and women (upper triangle)**

The highest correlation observed among men was with social Support (Pearson Correlation Coefficient 0.21), whereas, for female household head social cohesion presented the highest correlation with peer influence (Pearson Correlation Coefficient 0.20). The intracluster correlation (ICC) for the total sample was 0.04, indicating only 4% of the overall variability of tobacco consumption was attributed to the similarity within the community.
### Supplement 9: Table 6 - Sensitivity analysis – Comparison of multivariate odds ratios using “non-self cluster proportion” and “simple cluster proportion” of tobacco use among male household heads estimated by generalized estimating equation logistic regression in rural Uttar Pradesh, India (N= 5,312)

| Explanatory Variables | Adjusted model with non-self cluster proportion | Adjusted model with simple cluster proportion |
|-----------------------|-----------------------------------------------|---------------------------------------------|
| **Individual Demography** | | |
| Age categories (Ref: 18-29 Years) | | |
| 31-40 years | 1.34** (1.09, 1.63) | 1.35** (1.09, 1.65) |
| 41-50 years | 1.15 (0.95, 1.39) | 1.17 (0.96, 1.43) |
| 51-60 years | 0.96 (0.77, 1.19) | 0.99 (0.79, 1.23) |
| 61 years and above | 0.69** (0.54, 0.87) | 0.70** (0.55, 0.89) |
| Marital Status (Ref- Married) | | |
| Never married/Not stated | 0.75 (0.52, 1.09) | 0.74 (0.51, 1.07) |
| Widow/Separated | 0.91 (0.69, 1.20) | 0.92 (0.70, 1.22) |
| Caste (Ref- General) | | |
| ST/SC | 1.12 (0.95, 1.33) | 1.07 (0.91, 1.27) |
| OBC and others | 1.06 (0.90, 1.26) | 1.00 (0.84, 1.18) |
| Education (Ref- Illiterate) | | |
| Up to primary | 0.80* (0.68, 0.95) | 0.83* (0.70, 0.99) |
| Secondary | 0.57*** (0.47, 0.68) | 0.58*** (0.48, 0.69) |
| Above secondary | 0.30*** (0.24, 0.38) | 0.31*** (0.24, 0.39) |
| Occupation (Ref- Cultivator) | | |
| Wage laborer | 0.92 (0.79, 1.08) | 1.00 (0.85, 1.16) |
| Self-employed & Others | 0.98 (0.78, 1.23) | 1.02 (0.81, 1.29) |
| Unemployed | 0.78 (0.56, 1.08) | 0.77 (0.54, 1.10) |
| Household Wealth (Ref- Quintile 1) | | |
| Quintile 2 | 0.98 (0.79, 1.20) | 1.01 (0.81, 1.24) |
| Quintile 3 | 1.02 (0.81, 1.26) | 1.04 (0.83, 1.30) |
| Quintile 4 | 0.99 (0.78, 1.23) | 1.00 (0.80, 1.26) |
| Quintile 5 | 0.88 (0.69, 1.13) | 0.89 (0.70, 1.14) |
| Household Size (Ref- Small: up to 5 Member) | | |
| Large (>5 Members) | 1.02 (0.88, 1.16) | 1.00 (0.87, 1.16) |
| **Individual Personality Attributes** | | |
| Satisfaction with life circumstances (Ref- Low) | | |
| Medium | 1.02 (0.86, 1.21) | 1.00 (0.85, 1.18) |
| High | 0.95 (0.80, 1.14) | 0.92 (0.77, 1.09) |
| Level of happiness (Ref- Unhappy) | | |
| Neither happy nor unhappy | 0.78** (0.65, 0.94) | 0.80* (0.67, 0.96) |
| Happy | 0.82* (0.69, 0.97) | 0.85 (0.72, 1.01) |
| Perceived accessibility (Ref- Worsened) | | |
| Stayed the same | 1.57*** (1.32, 1.88) | 1.55*** (1.30, 1.85) |
| Improved | 1.66*** (1.35, 2.04) | 1.68*** (1.37, 2.05) |
| **Community Social capital** | | |
| Social Support | 1.02 (0.95, 1.08) | 0.91*** (0.87, 0.95) |
| Trust | 1.03 (0.97, 1.09) | 0.99 (0.95, 1.03) |
| **Peer influence** | | |
| Tobacco consumption in the community | | |
| Non-self cluster proportion | 1.10*** (1.05, 1.16) | |
| Simple cluster proportion | 1.58*** (1.53, 1.62) | |
| **Community demography** | | |
| Community wealth | 0.99 (0.92, 1.07) | 1.07* (1.01, 1.13) |
| Census blocks (Ref- Behadar) | | |
| Kachhauna | 0.96 (0.77, 1.18) | 0.96 (0.82, 1.11) |
| Kothwan | 1.08 (0.88, 1.31) | 0.98 (0.86, 1.11) |
| Kasmanda | 1.21 (0.96, 1.51) | 1.09 (0.94, 1.26) |
| Machhrehta | 0.91 (0.74, 1.12) | 1.01 (0.88, 1.15) |
| Sidhauli | 0.91 (0.75, 1.10) | 0.94 (0.82, 1.07) |
| Observations | 5312 | 5312 |

Note: AOR = Adjusted odds ratio, COR = Crude or unadjusted odds ratio, ST/SC = Scheduled castes and scheduled tribes, OBC = Other backward castes
**Supplement 10: Table 7 - Sensitivity analysis – Comparison of multivariate odds ratios using “non-self cluster proportion” and “simple cluster proportion” of tobacco use among female household heads estimated by generalized estimating equation logistic regression in rural Uttar Pradesh, India (N= 906)**

| Explanatory Variables | Adjusted model with non-self cluster proportion | Adjusted model with simple cluster proportion |
|-----------------------|-------------------------------------------------|----------------------------------------------|
|                        | AOR     | 95% CI          | AOR     | 95% CI          |
| **Individual Demography** |         |                 |         |                 |
| Religion (Ref- Hindu)   |         |                 |         |                 |
| Muslim and others      | 2.05**  | (1.21, 3.49)    | 2.16**  | (1.29, 3.63)    |
| Caste (Ref- General)   |         |                 |         |                 |
| ST/SC                 | 0.52*   | (0.31, 0.87)    | 0.52*   | (0.31, 0.88)    |
| OBC and others         | 0.75    | (0.45, 1.24)    | 0.75    | (0.45, 1.25)    |
| **Household Wealth (Ref- Quintile 1)** |         |                 |         |                 |
| Quintile 2             | 0.76    | (0.45, 1.30)    | 0.75    | (0.44, 1.29)    |
| Quintile 3             | 0.60    | (0.33, 1.09)    | 0.62    | (0.33, 1.14)    |
| Quintile 4             | 0.45*   | (0.24, 0.86)    | 0.49*   | (0.26, 0.90)    |
| Quintile 5             | 0.74    | (0.40, 1.37)    | 0.76    | (0.42, 1.38)    |
| **Household Size (Ref- Small: up to 5 Member)** |         |                 |         |                 |
| Large (>5 Members)     | 1.60*   | (1.04, 2.48)    | 1.55*   | (1.00, 2.40)    |
| **Individual Personality Attributes** |         |                 |         |                 |
| Satisfaction with life circumstances (Ref- Low) |         |                 |         |                 |
| Medium                 | 0.69    | (0.43, 1.08)    | 0.67    | (0.42, 1.07)    |
| High                   | 0.87    | (0.49, 1.53)    | 0.85    | (0.49, 1.45)    |
| **Individual social capital** |         |                 |         |                 |
| Community engagement   | 1.34**  | (1.08, 1.67)    | 1.32*   | (1.05, 1.66)    |
| **Community social capital** |         |                 |         |                 |
| Community engagement   | 1.18    | (0.93, 1.51)    | 1.24    | (0.98, 1.57)    |
| Social support          | 1.16    | (0.90, 1.51)    | 0.99    | (0.77, 1.28)    |
| Trust                  | 0.99    | (0.78, 1.26)    | 0.99    | (0.79, 1.24)    |
| Social cohesion         | 1.13    | (0.84, 1.51)    | 1.11    | (0.83, 1.47)    |
| **Peer influence** |         |                 |         |                 |
| Tobacco consumption in the community |         |                 |         |                 |
| Non-self cluster proportion | 0.97    | (0.85, 1.11)    |         |                 |
| Simple cluster proportions | 1.37*** | (1.17, 1.60)    |         |                 |
| **Community demography** |         |                 |         |                 |
| Community size (Ref- Small) |         |                 |         |                 |
| Medium                 | 1.23    | (0.78, 1.95)    | 1.31    | (0.84, 2.04)    |
| Large                  | 0.85    | (0.54, 1.36)    | 0.93    | (0.59, 1.48)    |

**Notes:**
- The estimates of the adjusted model with non-self cluster proportions will not match Table 4. Whereas, Table 4 did not include the non-self cluster proportion as it was not significant in the unadjusted model. However, in this sensitivity analysis, we have included the non-self cluster proportion variable in this model for comparing it with a simple cluster proportion.
- AOR = Adjusted odds ratio, COR = Crude or unadjusted odds ratio, ST/SC = Scheduled castes and scheduled tribes, OBC = Other backward castes.
### Supplement 11: Table 8 - Sensitivity analysis – Comparison between the odds ratio and prevalence ratio of tobacco use among male household heads estimated by generalized estimating equation rural Uttar Pradesh, India (N= 5,312)

| Explanatory Variables                          | Adjusted model with odds ratio |                | Adjusted model with prevalence ratio |                |
|-----------------------------------------------|-------------------------------|----------------|--------------------------------------|----------------|
| **Individual Demography**                     |                               |                |                                      |                |
| Age categories (Ref: 18-29 Years)              |                               |                |                                      |                |
| 31-40 years                                   | **1.34**<sup>**</sup>        | 1.09, 1.63    | **1.07**<sup>*</sup>                  | 1.01, 1.12     |
| 41-50 years                                   | 1.15                          | 0.95, 1.39    | 1.03                                 | 0.98, 1.09     |
| 51-60 years                                   | 0.96                          | 0.77, 1.19    | 0.99                                 | 0.93, 1.05     |
| 61 years and above                            | **0.69**<sup>**</sup>        | 0.54, 0.87    | **0.89**<sup>**</sup>                | 0.83, 0.96     |
| **Marital Status (Ref- Married)**             |                               |                |                                      |                |
| Never married/Not stated                      | 0.75                          | 0.52, 1.09    | 0.94                                 | 0.84, 1.05     |
| Widow/Separated                               | 0.91                          | 0.69, 1.20    | 0.96                                 | 0.89, 1.04     |
| **Caste (Ref- General)**                      |                               |                |                                      |                |
| ST/SC                                         | 1.12                          | 0.95, 1.33    | 1.04                                 | 0.99, 1.09     |
| OBC and others                                | 1.06                          | 0.90, 1.26    | 1.03                                 | 0.98, 1.08     |
| **Education (Ref- Illiterate)**               |                               |                |                                      |                |
| Up to primary                                 | **0.80**<sup>**</sup>        | 0.68, 0.95    | **0.95**<sup>*</sup>                  | 0.92, 0.99     |
| Secondary                                     | **0.57**<sup>***</sup>       | 0.47, 0.68    | **0.86**<sup>***</sup>               | 0.83, 0.90     |
| Above secondary                               | **0.30**<sup>***</sup>       | 0.24, 0.38    | **0.66**<sup>***</sup>               | 0.60, 0.73     |
| **Occupation (Ref- Cultivator)**              |                               |                |                                      |                |
| Wage laborer                                  | 0.92                          | 0.79, 1.08    | 0.98                                 | 0.94, 1.01     |
| Self-employed & Others                        | 0.98                          | 0.78, 1.23    | 0.99                                 | 0.93, 1.05     |
| Salaried worker                               | 0.78                          | 0.56, 1.08    | 0.92                                 | 0.82, 1.04     |
| Unemployed                                    | 0.80                          | 0.58, 1.09    | 0.95                                 | 0.86, 1.05     |
| **Household Wealth (Ref- Quintile 1)**        |                               |                |                                      |                |
| Quintile 2                                    | 0.97                          | 0.79, 1.20    | 0.99                                 | 0.94, 1.04     |
| Quintile 3                                    | 1.02                          | 0.81, 1.26    | 1.00                                 | 0.95, 1.06     |
| Quintile 4                                    | 0.99                          | 0.78, 1.23    | 0.99                                 | 0.94, 1.05     |
| Quintile 5                                    | 0.89                          | 0.69, 1.13    | 0.95                                 | 0.89, 1.02     |
| **Household Size (Ref- Small: up to 5 Member)**| 1.02                          | 0.88, 1.16    | 1.00                                 | 0.96, 1.03     |
| **Individual Personality Attributes**         |                               |                |                                      |                |
| Satisfaction with life circumstances (Ref- Low)|                               |                |                                      |                |
| Medium                                        | 1.02                          | 0.86, 1.21    | 1.01                                 | 0.97, 1.05     |
| High                                          | 0.95                          | 0.80, 1.14    | 0.99                                 | 0.94, 1.03     |
| Level of happiness (Ref- Unhappy)             |                               |                |                                      |                |
| Neither happy nor unhappy                     | 0.78**                        | 0.65, 0.94    | **0.94**<sup>**</sup>                | 0.90, 0.98     |
| Happy                                         | 0.82<sup>**</sup>            | 0.69, 0.97    | **0.95**<sup>*</sup>                  | 0.91, 0.99     |
| Perceived accessibility (Ref- Worsened)       |                               |                |                                      |                |
| Stayed the same                               | **1.57**<sup>***</sup>       | 1.32, 1.88    | **1.14**<sup>***</sup>               | 1.08, 1.21     |
| Improved                                      | **1.66**<sup>***</sup>       | 1.35, 2.04    | **1.16**<sup>***</sup>               | 1.09, 1.23     |
| **Community Social capital**                  |                               |                |                                      |                |
| Social Support                                | 1.02                          | 0.95, 1.08    | 1.00                                 | 0.99, 1.02     |
| Trust                                         | 1.03                          | 0.97, 1.09    | 1.01                                 | 0.99, 1.02     |
| **Peer influence**                            |                               |                |                                      |                |
| Tobacco consumption in the community          | **1.10**<sup>***</sup>       | 1.05, 1.16    | **1.03**<sup>***</sup>               | 1.01, 1.04     |
| **Community demography**                      |                               |                |                                      |                |
| Community wealth                              | 0.99                          | 0.92, 1.07    | 1.00                                 | 0.98, 1.02     |
| Census blocks (Ref- Behadar)                  |                               |                |                                      |                |
| Kachhauna                                     | 0.96                          | 0.77, 1.18    | 0.98                                 | 0.93, 1.04     |
| Kothwan                                       | 1.08                          | 0.88, 1.31    | 1.01                                 | 0.97, 1.07     |
| Kasmanda                                      | 1.21                          | 0.96, 1.51    | 1.04                                 | 0.99, 1.10     |
| Machhrehta                                    | 0.91                          | 0.74, 1.12    | 0.98                                 | 0.92, 1.04     |
| Sidhaut                                       | 0.91                          | 0.75, 1.10    | 0.98                                 | 0.93, 1.04     |
| **Observations**                              | 5312                          | 5312           |                                      |                |

Note: AOR = Adjusted odds ratio, APR = Adjusted prevalence ratio, ST/SC = Scheduled castes and scheduled tribes, OBC = Other backward castes.
### Supplement 12: Table 9 - Sensitivity analysis – Comparison between the odds ratio and prevalence ratio of tobacco use among female household heads estimated by generalized estimating equation rural Uttar Pradesh, India (N= 906)

| Explanatory Variables | Adjusted model with odds ratio | Adjusted model with prevalence ratio |
|-----------------------|-------------------------------|-------------------------------------|
|                       | AOR   | 95% CI       | APR   | 95% CI       |
| **Individual Demography** |       |               |       |               |
| Religion (Ref- Hindu) |       |               |       |               |
| Muslim and others     | 2.17**| (1.26, 3.72) | 1.85**| (1.18, 2.91) |
| Caste (Ref- General) |       |               |       |               |
| ST/SC                 | 0.53* | (0.31, 0.89) | 0.70  | (0.41, 1.17) |
| OBC and others        | 0.78  | (0.47, 1.30) | 0.95  | (0.62, 1.45) |
| Household Wealth (Ref- Quintile 1) |       |               |       |               |
| Quintile 2            | 0.75  | (0.44, 1.28) | 0.87  | (0.54, 1.40) |
| Quintile 3            | 0.57  | (0.31, 1.05) | 0.70  | (0.37, 1.31) |
| Quintile 4            | 0.43* | (0.23, 0.82) | 0.62  | (0.27, 1.42) |
| Quintile 5            | 0.70  | (0.37, 1.31) | 0.98  | (0.55, 1.74) |
| Household Size (Ref- Small: up to 5 Members) |       |               |       |               |
| Large (>5 Members)    | 1.60* | (1.03, 2.48) | 1.39  | (0.80, 2.41) |
| **Individual Personality Attributes** |       |               |       |               |
| Satisfaction with life circumstances (Ref- Low) |       |               |       |               |
| Medium                | 0.64  | (0.40, 1.01) | 0.71  | (0.44, 1.16) |
| High                  | 0.80  | (0.45, 1.42) | 0.78  | (0.40, 1.51) |
| **Individual social capital** |       |               |       |               |
| Community engagement  | 1.33* | (1.07, 1.66) | 1.34* | (1.03, 1.73) |
| **Community social capital** |       |               |       |               |
| Community engagement  | 1.18  | (0.93, 1.50) | 1.10  | (0.89, 1.37) |
| Social support        | 1.12  | (0.86, 1.46) | 1.22  | (0.94, 1.59) |
| Trust                 | 0.95  | (0.75, 1.22) | 0.86  | (0.66, 1.11) |
| Social cohesion       | 1.15  | (0.86, 1.54) | 1.09  | (0.82, 1.44) |
| **Community demography** |       |               |       |               |
| Community size (Ref- Small) |       |               |       |               |
| Medium                | 1.20  | (0.75, 1.92) | 1.26  | (0.75, 2.11) |
| Large                 | 0.89  | (0.56, 1.41) | 0.91  | (0.55, 1.49) |
| Observations          | 906   |               | 906   |               |

Note: AOR = Adjusted odds ratio, APR = Adjusted prevalence ratio, ST/SC = Scheduled castes and scheduled tribes, OBC = Other backward castes.