**Pre-analysis Plan for Smallholder Agricultural Competitiveness Project (SACP), Bangladesh**

Authors: Haseeb Ahmed, Silvio Daidone, Nicholas Sitko

# Background

The Smallholder Agriculture Competitiveness Project (SACP) accelerates implementation of the Master Plan for the Southern Agricultural Development by channeling investments to foster the competitiveness of the rural smallholders in the high value crops (HVC) and provides operational support to the Ministry of Agriculture (MoA). The SACP has three components: (1) enhance production of high-value crops (HVC) and technology adoption, (2) promote processing and marketing of HVC, and (3) strengthen climate resilient surface water management. The project implementation period is over six years and including the recently approved extension will be continued until June 2026. The impact evaluation was requested by the SACP team in year 3 of the implementation of the project and therefore we do not have a baseline before the start of the project.

The SACP supports smallholders to diversify their agricultural production systems away from predominantly rice production, to the production of other demand-driven HVC, with an associated intensification of production around homestead areas in Southern Bangladesh. The SACP supports 30 Upazilas from 11 Districts of this region in Bangladesh[[1]](#footnote-1). The region is susceptible to exogenous shocks, which have the greatest impact on poor people, including cyclones and storm surges, land erosion, limited livelihood opportunities, resource degradation, salinization, flooding, and waterlogged soils.

The total population in the project area is more than 7,000,000, representing more than 1,246,000 households. The project aims to directly benefit at least 250,000 rural households in the 30 Upazilas in Barisal, Chittagong and Khulna divisions. The selection of project beneficiaries was undertaken based on an inclusive targeting strategy focusing on small farmers.

The proposed study focuses on component 3 of the project: strengthen climate resilient surface water management. The activities under component 3 support households interested in increasing their productivity and diversification to HVCs with water infrastructure that provides supplemental or full season water access, through a range of investments in water storage and provision to cropland, with associated capacity development for households and groups to manage this water infrastructure. All activities in this component are closely associated and enhance activities under component 1 and 2 focused on HVC production and improved access to the markets. Activities included in component 3 are implemented by the Bangladesh Agriculture Development Corporation (BADC).

In the past, farmers in the SACP areas were limited to growing low-value crops due to the unpredictable nature of rain-fed agriculture. However, with the development of the surface irrigation schemes, farmers should be able to grow a wide range of high-value crops such as fruits, vegetables, and spices. These crops have higher market value and can be sold at premium prices, resulting in increased income for the farmers. Moreover, the availability of irrigation water should enable farmers to synchronize their crop production with market demand, including improved quality of the crops and reduce post-harvest losses. The project activities should facilitate formation of producers’ groups (or cooperatives) and other market-oriented initiatives that enable farmers to undertake collective and market-oriented actions.

# Theory of Change and Hypotheses Tested

This evaluation concerns three key water infrastructures built or rehabilitated by SACP. These infrastructures are:

1. Dykes along the rivers
2. Buried pipes
3. Canal rehabilitation and construction

Here an important distinction needs to be made between dykes and the rest of the water infrastructures. On the one hand, dykes protect the cropland from tidal waves and floods. By reducing flood risk, they allow for investment in agricultural production. On the other hand, buried pipes and canals increase the availability of water, especially during the Rabi season when rainfall precipitations decrease, thus allowing farmers to invest in agriculture. The evaluation design for dykes differs from the other water infrastructures. For the dykes, the evaluation design depends on household-level comparison of households that are along the river (within 1km distance), and are protected from the dyke, with households that are also located within 1 km distance of the river but are not protected from the dykes. However, for the other structures, the benefits accrue through water-user groups, and therefore the evaluation depends on clustered sampling. We discuss these topics in detail in Section 3. The hypotheses related to this evaluation are described below:

## 2.1 Hypotheses Regarding Crop Dykes

H1: Households protected from dykes will have greater overall resilience (measured through household food security).

H2: Households protected from dykes will face less flood risk.

H2: Farm production, land-used for agriculture, and investment in HVCs for beneficiary households will be higher because of enhancement in flood protection.

**Key Outcome Measures for the Dykes Evaluation**

* Food Security (FIES) – corresponding to H1.
* Self-reported flood shocks – corresponding to H2.
* Land under cultivation – corresponding to H3.
* Investment in HVCs - corresponding to H3.
* Revenue per decimal - corresponding to H3.

## 2.2 Hypotheses Regarding Canals & Buried Pipes

H1: Beneficiary households will have greater resilience and food security.

H2: Greater water availability, especially in the drier Rabi season, will allow for greater agricultural land use especially in Rabi season for beneficiary households as compared to the comparison group.

H3: Beneficiary households will have greater water availability, which will allow for agricultural diversification, high-value crop production, and increased farm production.

H4: Beneficiary households will have greater input use in terms of fertilizers and pesticides as they practice high-value agriculture.

**Key Outcome Measures for the Canals and Buried Pipe Evaluation**

* Food Security (FIES) – corresponding to H1.
* Land under cultivation – corresponding to H2.
* Number HVCs produced - corresponding to H3.
* Simpson’s index of agricultural diversification – corresponding to H3
* Revenue per decimal - corresponding to H3.
* Expenditure on agricultural inputs – corresponding to H4.

The hypotheses and outcomes are developed in light of the previous literature on the impacts of irrigation technologies on agricultural and household welfare outcomes. In the context of Rwanda, irrigation technologies allowed farmers to practice more high-value agricultural production (Jones et al. 2022). Similarly, Duflo and Pande (2007) show that better availability of water because of dam construction resulted in greater agricultural production and reduction in vulnerability to rainfall shocks in rural India. Several studies show that surface water irrigation schemes have a positive impact on agricultural productivity (see, e.g., Strobl and Strobl (2011), Zaveri et al. (2016)). Bravo-Ureta et al. (2020) find increases in the output of production (through greater agricultural land-use) but not in the technical efficiency of production as inputs factors do not change with the introduction of a canal rehabilitation project in Philippines. Based on these papers, we hypothesize that better water availability through irrigation infrastructures will increase high-value crop production, increase agricultural production through increased agricultural land-use and yield, and reduce the vulnerability to ‘less-than-normal’ (in the case of canals and buried pipes) and ‘greater-than-normal’ (in the case of dykes) rainfall in Bangladesh. Furthermore, because of increases in production and reduction in vulnerability to rainfall shocks, the food security of the beneficiaries will be strengthened.

# Study Design and Power Calculations

To evaluate the impacts of SACP infrastructural interventions, the impact evaluation is designed in two parts depending on the functionality of the structures involved in the study.

## 3.1 Part (a): The impact of Dykes on Households’ Socioeconomic Outcomes

To estimate the impact of dykes on household socioeconomic outcomes, we develop an impact evaluation design that relies on a treatment group, a transition group, and a comparison group. The definitions of these groups are provided below.

*Comparison group (C)*: Households along the river (within 1 km distance) where dykes will NOT be built.

*Transition group (T1)*: Households along the river (within 1 km distance) where dykes are yet to be built.

*Treatment group (T2)*: Households along the river (within 1 km distance) where dykes are already built.

At the baseline, the transition, and the comparison group both will form the overall comparison group. *T2* vs (*T1 + C*) will estimate the impact of dykes on high-value crop production, food security, and climatic resilience. Since there is no randomization between beneficiaries and non-beneficiaries of the intervention, we will rely on propensity score matching estimators and Inverse Probability Weighting Regression Adjustment (IPWRA) approach to estimate the impact of dykes on economic outcomes (Woolridge 2010).

But at the end line, comparing *T1* vs *C* will give the impact of dyke introduction on farmer wellbeing. We will use the staggered Difference-in-Difference design à la Callaway & Sant'Anna (2021). Such an impact evaluation design provides us the opportunity to evaluate the effectiveness of dykes through a cross-sectional comparison (in the first year of data collection) and then again allows us to validate these findings through a ‘phase-in’ DiD design (contingent on second year of data collection as well as *T1* getting the treatment), which is perhaps more powerful in establishing causal links between investment and outcomes.

In Table 1, we provide the sample size calculations based on Bangladesh Integrated Household Survey (BIHS) 2018-19 (IFPRI 2020). We create two indices of crop diversification from the BIHS data. The first is a Simpson index based on the share of income received from different crops (including vegetables and fruits). The second is the Standardized Weighted Average index based on the quantities of these crops harvested by the households. We present the sample size calculations and the associated minimum detectable effect sizes (MDES) in Table 1.

Procedure to estimate the sample sizes that will produce the desired MDES:

1. For our power calculations, we choose Simpson index of crop diversification based on share of income from different crops.[[2]](#footnote-2) The general formula for Simpson Index is:

where *shi* is the income share of crop *i*, calculated over total household crop income. We seek power of 80 percent[[3]](#footnote-3) and statistical significance of 5 percent.[[4]](#footnote-4)

1. We seek power of 80 percent[[5]](#footnote-5) and statistical significance of 5 percent.[[6]](#footnote-6) Because part of our evaluation has differing treatment timing and a longitudinal structure, it is important to take into account the autocorrelation structure of cluster-level errors (AR1). AR1 is set at 0.4, which is a commonly accepted level of autocorrelation in the impact evaluation literature.
2. The MDES is varied in the Schochet (2022) *Shiny* R app with assuming two treatment timings to estimate the number of observations required for the evaluation.
3. The MDES are presented in SDs as well as in % of the sample mean.

**Table 1: Sample Size Calculations for Evaluating the Impact of Dykes on Economic Outcomes based on Bangladesh Integrated Household Survey 2018-19**

|  |  |  |  |
| --- | --- | --- | --- |
| Outcome variable: Crop Diversification (Simpson index based on share of crop income from different crops) | | | |
| Sample Size per group | Total Sample Size | MDES (% increase in sample mean) | MDES (in SDs) |
| n = 243 | N = 729 | 25% | 0.25 SD |
| n = 332 | N = 996 | 21% | 0.20 SD |
| n = 459 | N = 1377 | 17.5% | 0.17 SD |
| Mean = 0.28; SD = 0.26, power = 0.8, AR1 = 0.4 | | | |

**Data Collection:**

* For this part, due to practical considerations, we will sample:
  + Comparison Group: 330 observations
  + Transition Group: 270 observations
  + Dyke Treatment group: 300 observations
  + Total observations: 900

## Part (b): The impact of Canals and Buried Pipes on Households’ Socioeconomic Outcomes

To estimate the impact of canals and buried pipe infrastructures on households’ socioeconomic outcomes, we develop an impact evaluation design that relies on a canal and buried pipes treatment groups, and a comparison group. The definitions of these groups are provided below:

*Comparison group*: Households in cropland parcels/villages where canals and buried pipe will NOT be built. However, these cropland parcels are within the same or adjacent mouzas[[7]](#footnote-7) where canals and buried pipes are built. Detailed data on the locations of canals and buried pipes was sourced from BADC and then mouza lists were developed in which these beneficiary water user groups existed. We also developed maps outlining the area where beneficiaries of canals or buried pipes could exist and comparison groups were chosen only outside these beneficiary areas, but within a proximate geographical location.

*Canal treatment group*: Households in WUGs/cropland parcels where canals are already built.

*Buried-pipe treatment group*: Households in WUGs/cropland parcels where buried-pipe are already built.

In Table 2 and 3, we provide the sample size calculations based on Bangladesh Integrated Household Survey (BIHS) 2018-19 (IFPRI 2020). The sample size calculations are again based on the Simpson Index and the Standardized Weighted Average index of crop diversification (Table 2 and Table 3, respectively).

Procedure to estimate the MDES for different number of comparison and treatment clusters and cluster sizes are given below:

1. For our power calculations, we choose two indicators of crop diversification that are of primary importance in this setting: (a) Simpson index of crop diversification based on share of income from different crops and (b) Standardized Weighted Average Index a la Anderson (2008) of crop diversification based on quantities of produce harvested.[[8]](#footnote-8) The general formula for Simpson Index is:

where *shi* is the income share of crop *i*, calculated over total household crop income. The standardized weighted average approach is based on Anderson (2008) and estimated using the ‘*swindex’* command in STATA (Schwab et al. 2020).

1. We seek power of 80 percent[[9]](#footnote-9) and statistical significance of 5 percent.[[10]](#footnote-10) Because part of our evaluation has differing treatment timing and a longitudinal structure, it is important to take into account the autocorrelation structure of cluster-level errors (AR1). AR1 is set at 0.4, which is a commonly accepted level of autocorrelation in the impact evaluation literature.
2. The intracluster correlation (ICC) is calculated using the ‘*loneway*’ STATA command.[[11]](#footnote-11)
3. The number of comparison and ‘transition’ group clusters are fixed at 30 each. Then number of treatment clusters and cluster sizes are varied in the Schochet (Number of treatment clusters information is fed into the Schochet (2022) *Shiny* R app to estimate the minimum detectable effect size.
4. The MDES are presented in SDs as well as in % of the sample mean.

**Table 2: Sample Size Calculations Based on Bangladesh Integrated Household Survey 2018-19**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Outcome variable: Crop Diversification (**Simpson index** based on shares of crop income from different crops) | | | | |
| Number of clusters per treatment arm (k) | Number of households per cluster (n) | Total Sample Size\* (N) | MDES in SDs | MDES (in % of sample mean) |
| 30 | 16 | N = 2400 | 0.31 SD | 28.4% |
| 30 | 15 | N = 2250 | 0.31 SD | 28.4% |
| 35 | 15 | N = 2475 | 0.29 SD | 26.6% |
| 35 | 14 | N = 2310 | 0.30 SD | 28.1% |
| 40 | 13 | N = 2340 | 0.29 SD | 26.6% |
| \*Assuming 30 parcels of pure comparison group and 30 clusters/parcels of “transition group” *for each iteration*. Transition group is where we expect the canals to be built in 2023/24.  Mean = 0.28; SD = 0.26, power = 0.8, = 0.05, ICC = 0.19, and AR1 = 0.4 | | | | |

**Table 3: Sample Size Calculations Based on Bangladesh Integrated Household Survey 2018-19**

|  |  |  |  |
| --- | --- | --- | --- |
| Outcome variable: Crop Production diversification (Standardized Weighted Average index based on share of total quantities harvested) | | | |
| Number of clusters per treatment arm (k) | Number of households per cluster (n) | Total Sample Size\* (N) | MDES (in SDs) |
| 30 | 16 | N = 2400 | 0.25 SD |
| 30 | 15 | N = 2250 | 0.25 SD |
| 35 | 15 | N = 2475 | 0.24 SD |
| 35 | 14 | N = 2310 | 0.24 SD |
| \*Assuming 30 clusters of pure comparison group and 30 clusters of “transition group” *in all iterations*. Transition group is where we expect the canals to be built in 2023/24.  Mean = 0; SD = 1, power = 0.8, ICC = 0.10, = 0.05, and AR1 = 0.4 | | | |

**Data Collection**:

*Taking into account the considerations from power calculations and practical considerations, the sampling was done as under:*

*Comparison group*: 40 cropland parcels/groups with 21 HHs within each parcel = 840 observations.

*Canal treatment group*: 40 WUGs with 19 HHs within each WUG = 760 observations.

*Buried-pipe treatment group*: 40 WUGs with 19 HHs within each WUG = 760 observations.

Total Observations = 2360 observations

# Estimation Strategy for Dykes Evaluation

In Year 1 of data collection:

We will estimate the following type of equation using the inverse-probability weighting approach (IPWRA) to balance for any observed, cross-sectional differences in the treatment and comparison groups:[[12]](#footnote-12)

Where are outcomes as outlined in section 2.1 for household *i* in mouza *j*. for households protected with a dyke (T2), while for households not protected with a dyke (T1 and C). is a vector of household characteristics to control for observable differences across households, which could influence Y. These factors are not only those for which some differences may be observed across treatment and control at the baseline, but also “good controls” which could have some explanatory role in the estimation of Y (Angrist & Pischke, 2009). are mouza fixed effects, while is the average treatment effect on the treated (ATET) of dykes on outcomes.

We also estimate the impact of dykes under a flood shock using the weighted difference-in-difference approach:

For outcomes such as food security and farm production, we expect and indicating that Dykes will increase the resilience of households in face of flood shocks. The variable *Flood shock* will be estimated using objective weather data. Here will be causal flood mitigation impact of dykes if *Flood Shock* is equally likely to occur across treatment and comparison groups.

In Year 2 of data collection:

A staggered difference-in-difference design is used to estimate the impact of dykes. The staggered design makes use of the phase-in period of the project and panel structure of the data to causally estimate the impact of dykes on economic outcomes and is a more robust causal inference model as compared to models based on cross-sectional comparison and matching.

The estimating equation is of the form:

Where is the outcome for household i, mouza *j*, and time *t*. for the households when they are protected by dykes in time , 0 otherwise. is our staggered DiD coefficient that reflects the causal impact of dykes on outcomes. are mouza fixed effects.

# Estimation Strategy for Canals and Buried Pipes Evaluation

We will estimate the following type of equation using the multivalued treatment effects inverse-probability weighting approach (IPWRA) to balance for any observed, cross-sectional differences in the treatment and comparison groups:[[13]](#footnote-13)

Where are outcomes as outlined in section 2.2 for household *i* in water user group (WUG) *j*. for canal beneficiary WUG, 0 otherwise. 1 for buried pipe beneficiary WUG, while and indicate comparison group. is a vector of household characteristics to control for observable differences across households, which could influence Y. These factors are not only those for which some differences may be observed across treatment and control at the baseline, but also “good controls” which could have some explanatory role in the estimation of Y (Angrist & Pischke, 2009). are group fixed effects, while and are the average treatment effect on the treated (ATET) for canals and buried pipes, respectively.

# References

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *Journal of the American statistical Association*, *103*(484), 1481-1495.

Angrist, J. D., & Pischke, J. S. (2009). Mostly harmless econometrics: An empiricist's companion. Princeton University Press.

Bravo-Ureta, B. E., Higgins, D., & Arslan, A. (2020). Irrigation infrastructure and farm productivity in the Philippines: A stochastic Meta-Frontier analysis. *World Development*, 135, 105073.

Callaway, B., & Sant’Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2), 200-230.

Cattaneo, M. D., Drukker, D. M., & Holland, A. D. (2013). Estimation of multivalued treatment effects under conditional independence. *The Stata Journal*, 13(3), 407-450.

Cattaneo, M. D. (2010). Efficient semiparametric estimation of multi-valued treatment effects under ignorability. *Journal of Econometrics*, 155(2), 138-154.

Duflo, E., & Pande, R. (2007). Dams. *The Quarterly Journal of Economics*, 122(2), 601-646.

International Food Policy Research Institute (IFPRI), 2020, "Bangladesh Integrated Household Survey (BIHS) 2018-2019", <https://doi.org/10.7910/DVN/NXKLZJ>

Jones, M., Kondylis, F., Loeser, J., & Magruder, J. (2022). Factor market failures and the adoption of irrigation in Rwanda. *American Economic Review*, 112(7), 2316-2352.

Schochet, P. Z. (2022). Statistical power for estimating treatment effects using difference-in-differences and comparative interrupted time series estimators with variation in treatment timing. *Journal of Educational and Behavioral Statistics*, *47*(4), 367-405.

Schwab, B., Janzen, S., Magnan, N. P., & Thompson, W. M. (2020). Constructing a summary index using the standardized inverse-covariance weighted average of indicators. *The Stata Journal*, *20*(4), 952-964.

Strobl, E., & Strobl, R. O. (2011). The distributional impact of large dams: Evidence from cropland productivity in Africa. *Journal of Development Economics*, 96(2), 432-450.

Wooldridge, J. M. 2010. Econometric analysis of cross section and panel data. MIT press

Zaveri, E., Grogan, D.S., Fisher-Vanden, K., Frolking, S., Lammers, R.B., Wrenn, D.H., Prusevich, A. and Nicholas, R.E., 2016. Invisible water, visible impact: groundwater use and Indian agriculture under climate change. *Environmental Research Letters*, 11(8), p.084005.

1. Upazillas are administrative divisions in Bangladesh, functioning as a sub-unit of a district. They correspond to a county or borough in Western countries. [↑](#footnote-ref-1)
2. These are not the only program outcomes that we are interested in. However, these are the most important outcomes and therefore rely on these outcomes for our power analysis. [↑](#footnote-ref-2)
3. This is the probability of correctly concluding that an intervention has an effect. It is commonly set at 80 percent in social sciences. [↑](#footnote-ref-3)
4. This is the probability of a false-positive result: the chance that a result shows that a treatment has an impact when in reality it does not. A broadly accepted threshold in the impact evaluation literature is 5 percent. [↑](#footnote-ref-4)
5. This is the probability of correctly concluding that an intervention has an effect. It is commonly set at 80 percent in social sciences. [↑](#footnote-ref-5)
6. This is the probability of a false-positive result: the chance that a result shows that a treatment has an impact when in reality it does not. A broadly accepted threshold in the impact evaluation literature is 5 percent. [↑](#footnote-ref-6)
7. A mouza represents an administrative unit, corresponding to a specific land area within which there may be one or more settlements. [↑](#footnote-ref-7)
8. These are not the only program outcomes that we are interested in. However, these are the most important outcomes and therefore rely on these outcomes for our power analysis. [↑](#footnote-ref-8)
9. This is the probability of correctly concluding that an intervention has an effect. It is commonly set at 80 percent in social sciences. [↑](#footnote-ref-9)
10. This is the probability of a false-positive result: the chance that a result shows that a treatment has an impact when in reality it does not. A broadly accepted threshold in the impact evaluation literature is 5 percent. [↑](#footnote-ref-10)
11. It is important to account for clustering when performing power calculations. The reason is that we expect the behaviours and hence the outcomes of beneficiaries (and non-beneficiaries) to be significantly correlated when they belong to the same cluster. This phenomenon is measured by the ICC: the higher the ICC, the lower the informational value of an extra observation from the same cluster. In other words, the ICC (that exists because of clustering) depreciates information, and this depreciation must be compensated for by either increasing the sample size, accepting a lower statistical precision, or considering a larger treatment effect size. [↑](#footnote-ref-11)
12. See, e.g., Woolridge (2010) for a detailed discussion of [↑](#footnote-ref-12)
13. See, e.g., Woolridge 2010; Cattaneo (2010); Cattaneo et al. (2013) for a detailed discussion on this methodological approach. [↑](#footnote-ref-13)