PIERI Project 20XXX

**Pre-analysis plan**

Insert the title of the report

*Insert the date*

**Abstract**

A summary of 12-16 lines (or between 100 and 150 words) briefly reporting the research background, main questions, results and policy implications of the study. Your abstract should reply these questions: What is known and why is this study needed. (Background and motivation); What you do? What are the main outcome variables and what is the intervention? (Methods); What do you find? (Results); Which are the policy implications? (Discussion)

**Key words** : Keyword 1, Keyword 2, etc.[[1]](#footnote-1)

**JEL Classification**: CODE1; CODE2; etc.

**Authors**

|  |  |
| --- | --- |
| **Author Name A**Professional titleInstitutionCity, CountryEmail address | **Author Name D**Professional titleInstitutionCity, CountryEmail address |
| **Author Name B**Professional titleInstitutionCity, CountryEmail address  | **Author Name E**Professional titleInstitutionCity, CountryEmail address |

**Instructions**

For an introduction to pre-analysis plans that will help you understand how to complete this template, please read the following entries:

* <https://blogs.worldbank.org/impactevaluations/a-pre-analysis-plan-checklist>
* <https://blogs.worldbank.org/impactevaluations/pre-results-review-journal-development-economics-lessons-learned-so-far>

This template is based on the *Journal of Development Economics* Pre-Results Review Stage 1 Submission Template. A reporting checklist table with research design details is available for download [here](https://docs.google.com/spreadsheets/d/1zhG3AxOJIB7H1m4-S0jwMN1cQKNVsLomsxktEDVrbP8/edit?usp=sharing).

Table of Contents

[I. Introduction 4](#_Toc97730406)

[1.1. Research question: background, importance and relevance 4](#_Toc97730407)

[II. Research Design 4](#_Toc97730408)

[2.1. Intervention(s) 4](#_Toc97730409)

[2.2. Hypotheses 4](#_Toc97730410)

[2.3. Timeline 5](#_Toc97730411)

[2.4. Basic methodological framework / Identification strategy 5](#_Toc97730412)

[2.4.1. Data 6](#_Toc97730413)

[2.4.2. Sample 6](#_Toc97730414)

[2.4.3. Data collection and processing 6](#_Toc97730415)

[2.4.4. Variations from the intended sample size 6](#_Toc97730416)

[2.4.5. Pilot data 6](#_Toc97730417)

[III. Empirical Analysis 7](#_Toc97730418)

[3.1. Statistical methods 7](#_Toc97730419)

[3.2. Statistical model 7](#_Toc97730420)

[3.3. Multiple outcome and multiple hypothesis testing 7](#_Toc97730421)

[3.4. Heterogeneous Effects 7](#_Toc97730422)

[IV. Baseline balance (First Interim Report instance) 7](#_Toc97730423)

[V. Progress of intervention (Second Interim Report instance) 8](#_Toc97730424)

[VI. List of Tables 8](#_Toc97730425)

[VII. List of Figures 8](#_Toc97730426)

[VIII. Appendices 8](#_Toc97730427)

[IX. Bibliography 8](#_Toc97730428)

# Introduction

## Research question: background, importance and relevance

* What is the main problem/question motivating the study? Why is this question important for the field of development economics?
* How has this problem/question been addressed thus far in the relevant literature? What are the competing theories for explanation of this question? How is this study different from prior research on this problem/question?

# Research Design

## Intervention(s)

* What type of an intervention does the study involve[[2]](#footnote-2)? Elaborate in detail when, where and by whom it will be delivered. Please provide sufficient detail to allow for replication in line with this journal’s [Mandatory Replication Policy](http://www.elsevier.com/__data/promis_misc/devec%20130805_ReplicationPolicy.docx).
* How will individual observations be assigned to treatment and control conditions[[3]](#footnote-3)?
* Are there multiple treatment arms involved and if so, are they exclusive or overlapping?
* What is the source of exogenous variation in your study?
* If applicable, what observations will be blinded (masked)[[4]](#footnote-4) after assignment to interventions and how? If blinding is not possible, what measures will be taken to minimize the potential for performance and expectancy biases (e.g. keeping participants unaware of trial hypotheses, measuring participant and provider expectations of benefit at baseline, etc.)?
* The instructions and supporting materials for the administration of the intervention should be included as an appendix.

## Hypotheses

* What are the main outcomes of interest? Which outcomes are primary to the analysis, which are secondary, and why?
* How will the main outcomes of interest be defined in your dataset?
* Please include all hypotheses which will be tested, linking each outcome specifically to how it will be measured. These should be reported as main results in the Stage 2 submission.

## Timeline

Please complete the table below with approximate dates for pending milestones in your research project. Feel free to add rows to make sure your mentor is kept informed of the upcoming calendar.

|  |  |  |  |
| --- | --- | --- | --- |
| **#** | **Milestone** | **START DATE** | **END DATE** |
| **1** | **Information campaign**Potential participants learn that the program exists and of its requisites |  |  |
| **2** | **Enrollment of potential participants**Potential participants register to be considered for the program |  |  |
| **3** | **Baseline data collection pilot and training of enumerators** |  |  |
| **4** | **Baseline data collection**Information is collected on potential candidates to learn about their state *before* the start of the program |  | No later than Aug 1st , 2022 as report is due Aug 31st, 2022 |
| **5** | **Randomization**Participants are randomly allocated to treatment and control groups. |  |  |
| **6** | **Intervention**Subjects in the treatment group undergo the program intervention.  |  | No later than February 28th, 2024 |
| **7** | **Baseline data collection pilot and training of enumerators** |  |  |
| **8** | **End line data collection**Information is collected on participants *after* the program ends |  | No later than March 31st, 2024, as report is due April 28th, 2024. |

## Basic methodological framework / Identification strategy

* What is the basic methodological framework of the study (RCT, pre-post, simple comparison, difference-in-difference etc.) and why is it suitable to address this research question?

### Data

Please use this section to provide details on *prospective* data that you will collect after acceptance of your research design.

### Sample

* What is the unit of analysis for this sample (individuals, organizations, etc.)?
* What is the expected sample size? If applicable, please include statistical power calculations[[5]](#footnote-5) to justify sample size.
* What is the effect size you will be able to detect?

### Data collection and processing

* What are the key data sources? What data collection procedures and instruments will be used?
* What is the rule for terminating data collection (number of observations, available funds, available time, etc.)?
* How long will the data collection process take? If data will be collected at multiple points (longitudinal design), what is the proposed schedule (including enrollment, intervention delivery and outcome assessment)?
* What measures will you take to ensure data quality in terms of accuracy, consistency, bias, and completeness (e.g. double entry if manual entry of handwritten forms, audits of survey work, etc.)?

### Variations from the intended sample size

* Do you anticipate any challenges in collecting data (attrition, non-compliance with the assigned treatment, etc.) and what measures do you plan to take to prevent them?

### Existing data

* Detail the sources for data used to establish reality checks, effect size estimations, feasibility, or proof of principle (ie. Basis for MDEs, ICCs, definition of outcomes, etc).

### Data storage

* Please state where you plan to store / are storing your data, in addition to storing a security copy after each collection stage in the PEP intranet, as required by the grant contract.

# Empirical Analysis

Please use this section to present your strategy for statistical analysis. In the appendices section of this submission, please also include any computer programs, configuration files, or scripts which will be used to run the experiment and to analyze the data.

## Statistical methods

* What statistical methods will be used to analyze the data and what are their underlying assumptions?
* How will the study deal with missing values?
* How do you define and handle outliers?

## Statistical model

Provide the model in its *functional* form and submit math equations as editable text and not as images.

## Multiple outcome and multiple hypothesis testing[[6]](#footnote-6)

* How will the study address false positives from multiple hypothesis testing?
* If you plan to adjust your standard errors, what adjustment procedure will you use? (e.g., Family Wise Error Rate, False Discovery Rates, etc.)
* If you plan to aggregate multiple variables into an index, which variables will you aggregate and how?

## Heterogeneous Effects

* Which groups do you anticipate will display heterogeneous effects? What leads you to anticipate these effects?

# Baseline balance (First Interim Report instance)

This section must present baseline summary statistics for the variables in the sample (obtained via survey or administrative sources). It must include:

* One or more baseline balance tables
* A brief discussion of balance results for key variables, stating why it might be important to control for some of this in the model specification, if applicable.

# Progress of intervention (Second Interim Report instance)

Use this section to provide detailed information on the progress of the intervention (date of roll out, coverage of geographical areas and experimental units, expected completion dates and possible setback if applicable).

# List of Tables

References can be in any style or format as long as the style is consistent.

# List of Figures

References can be in any style or format as long as the style is consistent.

# Appendices

If there is more than one appendix, they should be identified as A, B, etc. Formulae and equations in appendices should be given separate numbering: Eq. (A.1), Eq. (A.2), etc.; in a subsequent appendix, Eq. (B.1) and so on. Similarly for tables and figures: Table A.1; Fig. A.1, etc.

# Bibliography

Please respect the following format as it will simplify transition to the Working Paper template when the time comes.

Baah-Boateng, W. (2015). Unemployment in Africa: how appropriate is the global definition and measurement for policy purpose. International Journal of Manpower, 36(5):650–667.

1. Please consult the following link: <https://www.aeaweb.org/econlit/jelCodes.php>. By clicking on the JEL code links, you will find examples of keywords associated to specific domains. [↑](#footnote-ref-1)
2. For more information on reporting standards for interventions, see Hoffmann et al. (2014). [↑](#footnote-ref-2)
3. For more information on what to report on randomization, see Bruhn and McKenzie (2009). [↑](#footnote-ref-3)
4. Blinding or masking refers to methods of withholding information about assigned interventions post-randomization from those involved in the trial, when knowledge of this information could influence their behavior in a way that would later prove integral to interpreting the results(Grant 2017, 12) [↑](#footnote-ref-4)
5. Useful information and software tools for power calculations can be found [here](https://www.povertyactionlab.org/research-resources/software-and-tools). [↑](#footnote-ref-5)
6. Quality scientific publications will require that you think about this, so it is best to have it clearly laid out from the onset. If you are unsure on how to address this section, please reach out to your mentor. [↑](#footnote-ref-6)