Pre-analysis Plan for Cambodia ISAF 2019 - Analytical Core

Owen Ozier and Wei-Ting Yen*

June 18, 2019

Abstract

We examine an experiment undertaken by the Cambodian government, the World Bank, and an array of non-governmental organizations and bilateral donors. The intervention under study is the "Social Accountability Framework;" it emphasizes accountability for education, health, and commune services. The government of Cambodia permitted an experimental roll-out of this program: for the purposes of a randomized trial, the program took place in a randomly selected 21 of 42 possible districts, starting in 2017. The outcomes in this trial are measures of the quality of educational, health, and commune services provided to villagers in rural Cambodia, as well as measures of citizen engagement with the local government. Specifically, this trial tests whether a program that coordinates interventions by state and non-state actors can be successful in enforcing accountability and improving the quality of service delivery.

*Ozier: World Bank Development Research Group, IZA, and BREAD, oozier@worldbank.org. Yen: Franklin Marshall College, wyen@fandm.edu

1 Introduction

This document is an excerpt of a larger pre-analysis plan. Crucially, this piece of pre-analysis planning includes discussion of the construction of outcome measures, and the analytical strategy that will be used for those outcomes.

This portion of the pre-analysis plan was led by Owen Ozier, who joined the project in late 2018, with substantial input from Wei-Ting Yen; it incorporates details suggested by Kamakshi Mubarak and Erik Johnson.

This builds on an earlier trial registry entry, created in February 2017 and updated in March 2017. That registry entry was created by Wei-Ting Yen, and was written by Andrew Beath, Pham Trang, and Wei-Ting Yen.

That entry pre-dated the collection of baseline data. This document is written with access to baseline data gathered from March 13th to April 23rd in 2017.

This also includes an appendix with some of the content from the 2017 registry entry, to illustrate how closely the current analytical plans reflect the original intent.

This is written as endline data collection (April 18th -May 28th 2019) has just been concluded by the survey firm. However, those data have not been cleaned or shared with the authors of this document, nor has any analysis of the variables discussed in this document been shared with the authors of this document, or even undertaken by the survey firm, to the best of our knowledge. Thus we remain blind to any details of the endline data as we commit to this analysis.

2 Outcomes

The project team identified 32 indicators, grouped into five families, 31 of which are included in the main analysis. Each indicator is binary or ranges from zero to one. Most are measured based on survey interviews with residents of villages in the study area; some of these are restricted to respondents who are parents of children in primary school. The remaining

measures are derived from a survey visit to a health facility, school, or commune headquarters. All were gathered at baseline in early 2017, in addition to being included in the endline survey in 2019. This set of indicators closely follows the trial registry entry created in 2018, available at https://www.socialscienceregistry.org/trials/1995 .

1. Education
    - 1.1 Parent aware of maximum student-teacher ratio[1]
    - 1.2 Parent reports that child's teacher has not been absent during the past 30 days
    - 1.3 Parent reports that child received three free textbooks by December during the current school year
    - 1.4 Primary school has separate toilets for boys and girls with water for hand-washing
    - 1.5 Parent met with child's teacher this school year to discuss child's performance
2. Health
    - 2.1 Respondent reported that, during their most recent visit to the Commune Health Center (CHC) but within the past 12 months, the staff they encountered had been polite
    - 2.2 CHC has service fees displayed
    - 2.3 CHC has current year's budget information displayed
    - 2.4 Respondent reports that, during their most recent visit to the CHC but within the past 12 months, they saw the treatment fee listed
    - 2.5 Respondent is aware of the number of staff that should be present at the CHC during working hours
    - 2.6 Respondent is aware that the government requires the CHC to have medical staff available at night

[1] This indicator is defined as being within 3 of the correct answer: 42 is correct, but any answer between 39 and 45 (inclusive) is counted as correct for the purposes of this indicator.

- 2.7 Respondent reports that the number of staff actually working at the CHC during work hours is at least minimum
- 2.8 Respondent reports that, during their most recent visit to the CHC but within the past 12 months, they were informed of the treatment fee
- 2.9 Respondent reports that, during their most recent visit to the CHC but within the past 12 months, their consultation occurred in a location where other people could not hear their conversation with the medical staff
- 2.10 CHC has separate toilets for men and women
- 2.11 Respondent reports that, during their most recent visit to the CHC but within the past 12 months, the medical staff explained their condition and told them what they needed to do to make the condition better.
- 2.12 Respondent reports that medical staff are available at night

3. Council
    - 3.1 In the past 2 years, respondent applied and received certificate within three days[2]

4. Participation
    - 4.1 Respondent has attended a commune meeting in past year
    - 4.2 Respondent reports being aware of the commune budget for the previous year
    - 4.3 Respondent is aware of the commune projects implemented in their commune in the previous year
    - 4.4 Respondent participated in the commune project selection

[2]For any certificate that a respondent has requested (birth, death, or marriage), this is tallied. There can be more than one observation per respondent, since respondents could have applied for multiple certificates.

- 4.5 Respondent talked about commune, health or education related issues with local leaders outside of public meetings[3]
- 4.6 Fraction of council meetings that respondent attended in past year
- 4.7 Ratio of participants in last village meeting to the total number of households in the village

5. Voice
    - 5.1 Parent attended Parent Teacher Conference and discussed school issues
    - 5.2 Parent checked record book and left feedback for teacher this current school year
    - 5.3 **Not included in main analysis; see Section 3.2 below.** Respondent who experienced impolite treatment at the CHC during most recent visit within the past 12 months said something to the medical staff to change their behaviour[4]
    - 5.4 Respondent has put comments in a suggestion box or shared health-related concerns with Village Health Support Group members
    - 5.5 Respondent attended a commune meeting in the past year and spoke up during the meeting
    - 5.6 During the last village meting, residents shared their thoughts or opinions[5]

Clearly, some of the above measures are more easily checked (and less vulnerable to social desirability bias) than others. For measures regarding the facilities—whether fees are posted, functioning toilets for men and women, etc.—photo audits are used to verify the associated data.

[3]The specific list of relevant local leaders, for the purposes of service delivery on which this project focuses, as well as for the purposes of this indicator, was: any village chief, commune councillor, commune chief, community-based organization (CBO) staff, non-governmental organization (NGO) staff, monk, other religious leader, teacher, health centre staff, project management committee member, or village health support group (VHSG) member.

[4]The conditional nature of this indicator makes its interpretation potentially problematic. We therefore cut it from the main analysis, but include it in secondary analysis.

[5]Whether residents shared opinions is measured by asking village chiefs about the meeting.

In addition to these measures, we also ask a "process outcome" question, which we may refer to as our sixth measure. Though not included in the baseline survey, so we do not know its statistical properties, we ask respondents whether they have seen NGO staff in T-shirts in their community. In the event of a null finding on more important outcomes, this NGO staff present measure may be a useful diagnostic, providing some sense of the extent to which the program changed citizen exposure to NGO activities at all (a conceptual, if not actual, "first stage").

In treatment communities, we also expect to have detailed monitoring data from NGO implementers characterizing what aspects of the program were successful in the sense of being implemented according to plan. However, no comparable data will be available from comparison communities.

3 Main analysis

Our analysis strategy is as follows. Each of the 32 measures indexed by m, measured at the level of respondent or facility i in district d from measure family f, we can call $y_{i,d}^{f,m}$. We average all observations of a given measure within each district (for each of the 42 districts in the study) to construct $\overline{y}_{d}^{f,m}$. Further, within each district, all district-level mean measures in a given family can be averaged to form a single district-level composite index for that family, $\overline{y}_{d}^{f}$. The measures under the fourth and fifth families, "Participation" and "Voice," are intended to be disaggregated by respondent gender. Thus, there are seven families of hypotheses. Conceptually, then, a set of simple t-tests is desired for each of the seven associated null hypotheses.

A conceptually and econometrically straightforward way to approach this is to regress each of the five composite outcomes[6] on treatment in turn, estimating:

$$\overline{y}_{d}^{f} = \beta_1 Treatment_d + \epsilon_{1d} \tag{1}$$

The randomization was stratified on the basis of pairs: districts were

[6]The last two composite indices should be estimated separately by gender; this estimation strategy could also be used for the average NGO presence measure.

grouped into 21 pairs mainly on the basis of geography; one district from each pair was assigned to treatment, the other to control. Standard practice might be to include fixed effects for each pair P:

$$\overline{y}_d^f = \beta_2 Treatment_d + \sum_{p=1}^{P} \gamma_p + \epsilon_{2d} \tag{2}$$

Estimation of equation 2 might reduce statistical power, as the critical value of the t distribution will now be higher, though it might also improve statistical power if the pairs absorb much of the variation in the residual.

A potentially more powerful approach may be to include the baseline value of the composite outcome as a right-hand-side control variable. We do not know in advance the predictive power of this measure, but it is reasonable to consider including a term for $\overline{y}_{d,baseline}^f$ on the right side of either 2 or 1.

3.0.1 Process outcome analysis

The above analysis can also be performed for the "process outcome," in terms of whether any NGO activity is described by respondents (seen NGO officials in official shirts). A summary statistic in treatment areas of the fraction of respondents correctly identifying the color of NGO shirts in their area may also be informative.

3.0.2 Statistical robustness of main analysis

With this small number of observations, a final straightforward approach is to use randomization inference (RI) to obtain the distribution of the test statistic (the t-statistic or perhaps the coefficient from estimating either 2 or 1) under the sharp null by drawing from the 2^{21} possible randomized assignments to treatment, and testing against that distribution.

3.1 Disaggregated analysis: underlying indicators

If one of the families of indicators shows a statistically significant response to treatment, it may be informative to ask which underlying indicators appear to have driven that result. Those underlying indicators each have

multiple observations per cluster, since they are variously measured at the facility, parent, or respondent level. Inclusion of baseline values of the variables may also substantially improve statistical power. Thus we may estimate the following equation, clustering standard errors at the district level.

$$y_{i,d}^{f,m} = \beta_3 Treatment_d + \sum_{p=1}^{P} \gamma_p + \delta y_{i,d,baseline}^{f,m} + \epsilon_{3d} \tag{3}$$

From baseline data, we know that some of these outcomes have higher intra-class correlation than others. For example, the intra-class correlation of the individual-level measure of whether parents know the legal maximum for student-teacher ratios (indicator 1.1) is roughly 0.01; we have thousands of observations of this outcome, so precise estimation ought to be possible. However, whether field teams observed fees displayed at the health center (indicator 2.2) is much more correlated within district, with an intra-class correlation of approximately 0.21; we have fewer than 200 observations of this facility-level outcome. Thus, analysis of the former will be powered to detect much smaller changes than analysis of the latter.

3.1.1 Statistical robustness of disaggregated analysis

Regardless of power, with 42 clusters, estimation of Equation 3 may strike some readers as not being especially suitable for asymptotically motivated approaches to standard errors. Thus, the obvious robustness checks here are the wild cluster bootstrap of Cameron, Gelbach, and Miller (2011), and randomization inference (RI) adhering to the design of the randomization.

3.2 Secondary analysis: indicator robustness

Two variations on the construction of indicators, and any associated changes to the analysis–both regarding levels of, and impacts on, aggregate family-level composites–are anticipated to be important to report for completeness, though they should not be considered primary analyses from a statistical perspective.

First, one of the indicators, 1.1, measures whether adults know the maximum legal student-teacher ratio in primary school. The final construction of this indicator for primary analysis is restricted to parents of children in school, and allows any answer between 39 and 45 to be considered correct. The secondary analysis is to re-create the original indicator as described in early project documents: it is asked of all adults regardless of whether they have school-age children, and only the exact answer 42 is accepted as correct.

Second, another of the indicators, 5.3, measures whether respondents who experienced what they judged to be impolite treatment at the health facility then said something to medical staff in response. This is clearly conditional on the presence of impolite behavior, so if that rate changes, this indicator becomes difficult to interpret. It is therefore excluded from family-level composite measure construction in primary analysis, but because it was described in early project documents, it should be included in family-level composite measure construction in secondary analysis.

3.3 Secondary analysis: heterogeneity

Though it may be less statistically powered, we expect secondary analysis to report program impacts separately by ethnicity. The most natural split in relation to the study's topic–political empowerment–might be to consider Khmer (majority) indigenous groups as one "ethnicity," and to group together all non-indigenous national minority groups (Vietnamese, Lao, Thai) as another. Among groups indigenous to Cambodia, it may also be helpful to conduct impact analysis separately for Cham/Muslim respondents, separately from Khmer respondents; we anticipate Cham/Muslim groups to be the majority of households, after Khmer.

A Original Trial Registry Appendix

Though we do not intend to precisely follow the notes below, we include them for completeness to show how thinking has evolved in this project from 2017 to 2019. At that time, the project team had not committed to

outcomes, even internally within the World Bank. There is nevertheless substantial overlap between what had been noted in the registry entry, and our current analytical plans.

A.1 Excerpt from Trial Registry in March, 2017

A.1.1 Primary Outcomes (end points)

Key outcomes include: School Enrolment, Attendance and Repetition; Test Scores; School Services and Facilities; Satisfaction with Education Quality; Unofficial Cost of Education; School Performance; School Engagement; School Awareness; Infant Mortality; Utilization of Health services; Health Services and Facilities; Access of Poor Villagers to Health Services; Cost of Healthcare; Use of Non-Professional Health Providers; Health Service Performance; Satisfaction with Health Services; Health Service Engagement; Ownership of Birth Certificates; Quality of Service Provision; Unofficial Cost of Service Provision; Access to Information; Satisfaction with Service Provision; Commune Council Performance; Attendance in Village or Commune Meetings; Engagement with Leaders; Social Capital; Awareness; Use of Information

A.1.2 Primary Outcomes (explanation)

Education Outcomes: Enrolment; Attendance; Repetition; Test Scores: Successful completion of basic numeracy and literacy test; School Services and Facilities: Audit of school facilities (teacher and staff absenteeism, desks, chairs, textbooks, bathroom facilities, electricity, blackboards / whiteboards etc.); Unofficial Cost of Education: Incidence and amount of unofficial payments for textbooks; Incidence and amount of unofficial payments for (involuntary) private tutoring; Incidence and amount of unofficial payments for enrollment.; School Performance: Attainment of government performance standards; School Engagement: Frequency of Discussions with School Teachers and Officials; School Awareness: Parental attendance of teacher-parents meetings; Parental access to student record book; Parental awareness of presence of school support committee; Parental

awareness of free textbook policy; Parental Awareness of School Budget; Health Outcomes: Infant Mortality; Health Services and Facilities: Audit of health facilities (staff absenteeism, availability of basic medicines; cleanliness of facilities; 24 hour attendance; number of patients waiting and duration of wait).; Access of Poor Villagers to Health Services: Incidence of ownership of HEF cards among poor villagers; Awareness of health care entitlements among IDpoor households.; Cost of Healthcare: Total unofficial payments for health services in the past year; Proportion of unofficial payment over total medical expense in the past year; Total medical expenses in the past year.; Use of Non-Professional Health Providers: Frequency of self-treatment in event of illness; Frequency of visits to traditional healers in event of illness; Frequency of home delivery without qualified midwife.; Health Service Performance: Attainment of government performance standards; Health Service Engagement: Frequency of discussions with health service administrators; Service Outcomes: Ownership of Birth Certificates for Children Aged 1 - 4; Unofficial Cost of Service Provision: Frequency of payment of unofficial fees for certificates; Access to Information: Availability of documentation on commune budget; Perceptions of trustworthiness of published documentation of commune budget; Perceptions of trustworthiness of published documentation of local development projects.; Satisfaction with Service Provision: Satisfaction with work of village chief and commune councilors in past year; Perceived benevolence of village chiefs; Perceived benevolence of commune councilors; Perceived benevolence of district councilors. ; Commune Council Performance: Attainment of government performance standards; Meeting Participation: Participation in Village or Commune Meetings ; Social Capital: Perceived benevolence of other villagers; Acquaintance with villagers other than relatives; Acquaintance with villagers of marginalized status; Interpersonal trust; Incidence of socially-cooperative behavior.; Awareness: Awareness of council decisions and budgetary allocations; Awareness of meeting times and topics of last council meeting.; Use of Information: Knowledge and understanding of information and budgets.