Nepal Health Insurance Pilot
SIEF Technical Proposal

I. Background and Overview

Healthcare in Nepal

Nepal is country of 23 million, half of whom live below the poverty line. The country has seen significant improvements in health indicators over the past 20 years, but lags behind its neighbors on key indicators such as infant and maternal mortality, childhood malnutrition, and life expectancy. Furthermore, these indicators vary greatly between regions and income levels: life expectancy in in the capital of Kathmandu is 74 years; in the mountains of Mugu district it is half that.[1] Nepal has a wide network of public health facilities, but due to distance and cost only about 50% of the poorest Nepalese seek care when they are ill. The average treatment cost per illness episode is equivalent to 1.5 months of per capita consumption for the very poor.[2] Bringing this cost of use down is necessary to increase access to health services by the poor and other marginalized groups, including women.

In 2007, the Government of Nepal declared that essential health care services (EHCS) would be delivered to all of its citizens free of cost in an effort to improve the accessibility to basic health services for the poor. Under EHCS, the government has identified 20 key primary care service interventions to subsidize including those relating to safe motherhood, family planning, outpatient care, and child health.[3] In principle, EHCS also covers many of the most common drugs; however, the demand for these drugs often outstrips supply at state-run health facilities, resulting in rationing and the need for patients to purchase additional drugs from private facilities.

Despite the provision of free EHCS, public expenditure on health is currently low: only 7% of the budget or US$10 per capita. Signaling a desire to increase this level of expenditure, the Ministry of Health and Population (MoHP) has been working for the past year to develop a national health financing strategy in order to realize Nepal's vision of universal coverage. A parallel effort is ongoing to pilot a health insurance that aims to address the two most important challenges facing the sector: rising out of pocket expenditures and increasing inequities in access to and utilization of health care across income quintiles and geographical regions.

Out-of-pocket (OOP) expenditures incurred by Nepali users of public hospital services are approximately US$34 per month[4], an amount that approaches the average total household consumption for most

[1] Nepal Health Sector Program Appraisal Document (PAD). World Bank. 2004.
[2] Country Assistance Strategy for Nepal. World Bank. 1998.
[3] Effectiveness of Essential Health Care Services Delivery in Nepal. Nepal Health Resource Council. 2008.
[4] Note that this figure is only for households who used hospitals services that month- This figure does not represent average monthly hospital expenditures for the year since most families will not continue to use hospital services on a monthly basis.

regions.[5] Total OOP (for both primary and secondary care) associated with the use of publicly provided health care represents around 40% of total monthly household consumption for middle quintiles, significantly less for the top quintile (<30%) and significantly more for the poorest quintile (>50%). In part due to health care's prohibitively high cost relative to total consumption, the poorest households don't use health services as much as wealthier ones: the utilization of public hospital services for the poorest quintile is half that of the second poorest quintile and less than a third that of the second wealthiest quintile.[6] This gap in health care utilization is almost certainly understated: wealthier households tend to use more expensive private facilities that are not captured in these figures.

Universal Health Coverage

The concept of universal health coverage to address the inequities in health care access between rich and poor has grown in visibility and importance in recent years. In 2010, World Health Organization member states adopted a resolution encouraging countries to transition to universal coverage. Despite efforts at universal coverage in a number of mid- and low-income countries, there is little available evidence on the causal links between the design features of universal health insurance and key outcomes.[7] The available evidence overwhelming conforms that health insurance schemes improve access and utilization on average, but it is not clear on the relative benefit to the poor and other vulnerable groups.[8] Most studies find an effect on out of pocket expenditures despite two underlying drivers working in opposite directions: price-per-visit is often lower, but since utilization is increased, the drop in total health expenditures may not be as great.[9] Finally, the overall evidence of the impact of health insurance on health status is inconclusive; very few studies have attempted to estimate the causal link between insurance coverage and health outcomes and many that have tried suffer from "unresolved methodological challenges and important study limitations."[10]

There is reason to believe that these health insurance schemes that aim to achieve universal coverage have a differential effect on the poor. A number of evaluations have been conducted of tax-payer funded insurance schemes (as opposed to schemes that rely on premiums paid by beneficiaries), most of which are targeted to the poor. These studies suggest that free insurance to the poor tends to increase health care utilization, but does not reduce out-of-pocket expenditures (see China[11] and Vietnam[12]). Furthermore, for the poorest decile of the population, the increase in utilization is negligible.

A difficulty for many of these studies is that the take-up rates for even highly subsidized voluntary insurance schemes tend to be low: In Nicaragua[13], take-up of a subsidized insurance program for

[5] Benefit Incidence Analysis (BIA) in Health – Nepal. Oxford Policy Management. 2012.
[6] BIA. Pages iii-13.
[7] The Impact of Universal Coverage (UC) Schemes in the Developing World: A Review of the Existing Evidence. World Bank. 2012.
[8] The Impact of UC Schemes. Page 50.
[9] The Impact of UC Schemes. Page 60.
[10] The Impact of UC Schemes. Page 68.
[11] Wagstaff, et al. *Extending Health Insurance to the Rural Population: An Impact Evaluation of China's New Cooperative Medical Scheme.* 2009.
[12] Wagstaff. *Health Insurance for the Poor: Initial Impacts of Vietnam's Health Care Fund for the Poor.* 2007.
[13] Thornton, et al. *Social Security Health Insurance for Informal Sector in Nicaragua: Randomized Evaluation*. 2010.

informal sector workers was around 20% with less than 10% still enrolled after one year. In the state of Karnataka in India[14], take-up of a catastrophic health insurance policy was higher, but only after it was made mandatory for individuals interested in borrowing from a microfinance institution. In Mexico[15], where the Segura Popular insurance is free for the poor, take-up rates were much higher (44%), though coverage is still nowhere near universal. Recent experience from India shows that less than 50% of the below the poverty line population (eligible for the seminal public health insurance program (RSBY)) is actually enrolled in RSBY. The team hopes to address this issue in Nepal with an encouragement design to improve take-up.

The findings intended to be generated through this pilot are directly related to the evidence gaps outlined in the Strategic Impact Evaluation Fund (SIEF) cluster note. In particular, the cluster note points out that there is little evidence on how to stimulate demand for health services and promote health from within communities, which this pilot will examine. While a primary concern of the pilot is to inform the Government of Nepal's plans for the national scale-up of the health insurance pilot, this impact evaluation will also contribute to the literature on the impact of health insurance schemes on health outcomes and the impact of encouragement mechanisms on take-up (with which many of the studies that make up the current literature struggle).

II. Objectives

The primary objectives of the impact evaluation are to:

1. Evaluate the effectiveness of the health insurance pilot for the at-risk population in:
    a. Increasing access to and utilization of health services;
    b. Reducing out of pocket expenditures; and
    c. Improving drug availability

2. Conduct a process evaluation of the health insurance to illuminate potential constraints and incentive problems to inform the national scale up. The process evaluation will also contribute to explaining the impact evaluation results. The process evaluation will have three components:
    a. *Administrative.* Evaluation of the quality of monitoring carried out by the health insurance, and the ability of the health insurance to manage insurance-related risks such as adverse selection, over-consumption and over-prescription.
    b. *Operational.* Evaluation of the ability of the health insurance to respect the basic insurance principles such as to provide benefits in return for premiums.
    c. *Financial.* Evaluation of the ability of the health insurance to rely on its revenue to cover expenses over the long term.

[14] Chen, Comfort, and Bau. *Implementing Health Insurance through Micro-Credit: A Case Study of SKS Microfinance, India.* 2008.
[15] King, et al. *Public Policy for the Poor? Randomised Assessment of Mexican Universal Health Insurance Programme.* 2009.

3. Provide cost analysis data to inform decisions relating to coverage limits and price negotiations and to estimate the cost-effectiveness of results for the national scale-up.

III. Description of the Intervention

Insurance Beneficiaries

All members of an insured household will be covered. In case of non-resident members, a minimum of six months residence in the previous year is required to qualify for coverage. The ultimate goal is to cover all Nepalese in the selected districts and eventually throughout the country. However, given the limited resources the coverage must expand gradually. Gradual expansion means that choices have to be made as to which population geographic and group should be given priority. First the five districts were selected to limit the geographic coverage of the pilot. The insurance product will be made available to resident in the selected districts. Second priority was given for the poor in terms of spending public resources so that the premium of the poor will be subsidized. This would require identifying households to be targeted. Recognizing the challenge in identifying households, the pilot will use proxy means testing as a targeting tool to identify households. This tool has been used for targeting various social programs in a number of countries, including Nepal. The pilot will take into account the role of VDC officials in supporting the process of identifying households within the selected VDC.

Once households are identified, a nominal enrollment fee will be assessed per household. The enrollment fee is meant to enhance households' sense of ownership of the scheme. The program administrator will be incentivized through a payment system that encourages enrolling as many households in the targeted districts as possible. These efforts will also be supplemented by marketing of the scheme.

Benefits Package

In defining the benefit package, it is important to take into account the context of Nepal where certain services and medicines are already provided free of charge to the user (under the existing EHCS). A conscious and deliberate decision was made to leave out services that are part of the free care from the benefit package of this pilot insurance product. However, there is growing evidence showing that drug and medication constitutes a significant proportion of out of pocket payments in Nepal. The irony is that a significant portion of these drugs are part of the free care. This seemingly inconsistent fact is explained partly because of drug stock outs at the health facilities. Even if the drugs are free at the public health facility, when the facilities run out of stock, patients are required to purchase the drugs from outside the facility. The benefit package therefore takes into account the following: (i) drugs that are meant to be free through EHCS but are not always available; (ii) there are services that are not free through EHCS that households continue to pay out of pocket; and (iii) not all segments of the Nepali society have access to free care.

Based on these considerations, the benefit package includes all drugs that are part of the free care and all in-patient services. However, the benefits an individual can receive annually will be capped to ensure the sustainability of the scheme. The drug benefit and in-patient service benefits will have separate caps of

different amounts. As such beneficiaries will have guaranteed services for in-patient and drugs/medicines up to a certain amount. This approach has been chosen over an alternative where beneficiaries select the coverage-copay mix. The list of drugs covered is attached in Appendix B.

Service Providers

Beneficiaries will choose the provider and obtain services covered by the insurance from either public or private providers. However, the scheme will approve providers; beneficiaries will be given the list of providers approved by the scheme that they can choose from. The approval takes the form of pre-qualification by the scheme. The process will allow the scheme to negotiate prices for services covered by the scheme, place minimum quality requirements for services, etc. Once the scheme qualify providers, the list of providers in each district approved by the scheme will be published.

Insurance Administrator

Given the level of sophistication from all parties, the scheme will start with a less complex arrangement where an administrator (as opposed to an insurer[16]) will be hired on a competitive basis to administer the insurance scheme. The bidders for the administrator role will compete on both prices (premiums that members have to pay), and caps on annual coverage. Payment to the administrator will be based on the number of beneficiaries enrolled. The administrator is responsible for processing claims and making payment to providers. It also will work with the MOHP in developing criteria for qualifying providers.

Premiums and Payments to Provider

The administrator will be responsible for collecting premiums and paying for service providers. Payments will be made electronically where each member (household or individual) will have an e-card with which payments can be made. Electronic payments are critical to reduce the need for the patient to make payments at the time of service utilization and to render the system cashless.

Results Chain

The **inputs** to the intervention include the current health care infrastructure including hospitals and clinics, health care workers, and the existing coverage scheme EHCS (Essential Health Care Services). The key inputs that will be added to this current system include a new subsidized *insurance product* and an *administrator* (to be hired on a competitive basis) who will be responsible for enrolling participants, processing claims, and making payments to providers.

The following are the primary **activities** that make up the intervention:

- Hire administrator who will be paid on the basis of number of beneficiaries enrolled
- Design insurance product (premium levels, coverage details and limits, payment/reimbursement mechanisms)
- Qualify health service providers to be included in the scheme
- Negotiate prices for services with providers

[16] The distinction between administrator and insurer is the extent of risk to borne by the party. If it is administrator, no risk is to be borne by it and a fixed fee will be paid for providing the services in administering the scheme. On the other hand, if the party is to assume some risk it operates as an insurer where it can make loss/profit from the premium it collects and the payments it makes to service providers.

- Enroll participants with marketing and additional encouragement mechanisms
- Facilitate payment to providers

The immediate **outputs** of these activities will include the insurance product itself and, assuming effective encouragement mechanisms, a reasonable rate of take up among the households in the targeted districts. Other outputs include the number of total beneficiaries, ID cards produced (for payments), and healthcare providers qualified. Based on the studies cited in the background section of this document, the team expects for a high degree of enrollment in the program to lead to the following **outcomes** (consistent with the outcomes emphasized in the health cluster note):

- Increased access to and utilization of health care services
- Decreased out-of-pocket expenditures for health services

IV. Research Questions and Policy Relevance

Research Questions

The primary research questions for this impact evaluation are:

1. What is the impact of the subsidized health insurance intervention on:
   - Access to and utilization of health services?
   - Out-of-pocket payments for health services?
   - Improvements in drug availability
2. What factors determine the take-up of the insurance?
3. What institutional factors are associated with the outcomes of interest?

We will exploit the information of premium subsidy that targets poor households, to identify the impact of the insurance with subsidy.

Furthermore, all of the research questions will be analyzed for heterogeneous impacts based on income quintile and disadvantaged groups.

Policy Relevance

The results of this evaluation will be used to inform the design and implementation of Nepal's national health coverage plan. The impact evaluation will answer important questions related to the effectiveness of the insurance scheme especially as they relate to key vulnerable groups, including the poor and minority groups. Further, the cost analysis and process evaluation will provide critical operations data that will be immediately relevant to the Government of Nepal as it finalizes and rolls out the next phase of its national health care plan.

V. Evaluation Design

This evaluation consists of three treatment districts () where the new health insurance scheme will be rolled out. Within these five districts are 153 Village Development Committees (VDCs).[17] The impact evaluation of provision of new health insurance scheme will be primarily based on a randomized enrolment promotion/encouragement design targeted to increase take up rate of the new health insurance product. However, since the feasibility of this method in identifying the impact is highly dependent on the efficacy of the randomized encouragement on the take up of health insurance, a provision also is made to incorporate a quasi-experimental design which could be utilized for assessing the impact in case the randomized encouragement design does not result in sufficient take up differential. Specifically, under such conditions, a matched difference-in-difference design will be used for impact evaluation. Thus, the evaluation will combine both experimental as well as quasi-experimental designs, where the experimental design will be given priority over the quasi-experimental design.

A. Designing encouragement mechanism for increasing take-up rates

It is envisioned that the MoPH will identify the targeted beneficiaries in these three pilot districts and undertake appropriate strategies to enroll the beneficiaries.

As described in the background section, take-up rates for insurance schemes in low-income countries tend to be low. Both the success of the program (the Government of Nepal is interested in universal coverage) and the success of the evaluation depend on take up rates.

In order to ensure sufficient take-up, the team will use an enrolment promotion/encouragement strategy which will act as a booster to the regular enrolment process.

The enrollment encouragement intervention will include various events at the community level including meeting with the community opinion leaders, undertaking information dissemination through schools, undertaking mass enrollment drive in the VDC and additional follow up events such as household visits for select households. The information provided will be limited to the enrollment and not about use of health services or similar information that would change the information set related to the decision to use services that households would have in the absence of the encouragement.

The 153 VDCs in the 3 districts will be randomly assigned to treatment and control group where the treatment VDCs will receive enrollment encouragement intervention. Approximately half of these VDCs will be randomly assigned to treatment group that will receive the promotion and other half will be assigned to control group which will not receive any such promotion.

The differences in take-up rates between the treatment and control VDCs is the estimates of the effect of the encouragement mechanisms on the take-up rate.

[17] VDCs are similar to municipalities in other countries, but with greater public-government interaction. Source: http://en.wikipedia.org/wiki/Village_development_committee

Table 1: Encouragement Intervention

Encouragement Mechanism	Community+ Household Registration (Treatment)	No Additional Promotion (Control)
No. of VDCs	77	76

If this effect is sufficiently high – that is, if VDCs receiving the promotion are significantly more likely to take-up the insurance program than those that did not receive the promotion – the team will be able to use the randomized-promotion design to estimate the impact of the program on the relevant outcomes by utilizing the encouragement assignment as instrument of participation. However, if the take-up differential is not sufficiently high, the randomized-promotion design will not be useful. In this case, the team will use households in neighboring districts to allow for a matched difference-in-difference design which we discussed later.

Preliminary estimates of the trade-off between standardized effect size and sample size (for standard power needs) indicate that the encouragement design needs to generate a take up differential in the range of 25-35% between the treatment and control groups in order to be able to detect a reasonable effect size of health insurance on utilization of access to health care facility. [18]

The section below describes the strategies to be adopted for the evaluation of the program.

B. Estimating the impact of the health insurance program on access, out of pocket expenses, utilization, and health outcomes (all 3 districts).

Experimental design

As explained above, if the encouragement mechanism is found to be effective in increasing take-up rates (ensuring minimum 25% take up differential), the team will use the encouragement mechanism to estimate the impact of the health insurance program on the outcomes of interest using VDCs in all three target districts. This will be done by taking advantage of the differential take-up rates of the groups which will receive the encouragement treatment and those which won't in a **randomized-promotion design**. The encouragement assignment (which will be randomized) will be used as an instrument for participation in the program (which will not be randomized- all households in the three districts will be offered the program and can self-select to participate).[19] Assuming the encouragement mechanism results in a significantly higher take-up rate in the treatment sample as compared to the control sample of the encouragement treatment (generating a take up differential of 25% or more), we can conduct a two-stage least squares regression where we isolate the variation in participation rates caused by the encouragement

[18] Please refer to the note in the appendix for the details on the success rate calculation

[19] Assignment to the encouragement mechanism is an appropriate instrument for participation in the program as this assignment (a) is correlated with participation in the program (assuming the encouragement increases take-up rates); and (b) is not correlated with the outcomes of interest except through its effect on participation in the program.

mechanism and use that to estimate the differences in overall outcomes (utilization and out of pocket expenditures). These differences can be interpreted as the *local average treatment effect*- in other words, the effect of the health insurance program on those households who would be affected by the encouragement mechanism.

Table 2: VDC Assignment to Randomized Promotion and Control Groups by District

District	Ilam	Baglung	Kailali	TOTAL
Encouragement	25	30	22	77
No Encouragement (Control)	24	30	22	76

Quasi- experimental design

In the unlikely event that the encouragement mechanism does not significantly increase program take-up rates, the team will estimate the impact of the health insurance program on the outcomes of interest with a matched difference-in-difference design.

We propose using a matching method where in the first stage, 3 districts will be selected as control districts which will be comparable to the treatment districts with respect to important parameters (such as population, health infrastructure, utilization of healthcare etc.). In the next stage, a VDC level survey will be undertaken to collect VDC level data on important parameters that will be used for matching. Using the VDC level data VDCs in the three treatment districts will be matched with a VDC of similar characteristics in the neighboring (non-treatment) districts (to be referred to as *pseudo-control* VDCs).[20] To ensure that adequate matches are identified, we will oversample VDCs from surrounding districts that are comparable to the treatment districts.

The following section describes the detailed matching strategy to be adopted to identify the *pseudo-control* VDCs.

Stage1: Identification of matching districts

In the first stage district level secondary data will be utilized to identify 3 districts with similar characteristics as those of the treatment districts and the pseudo- control VDCs will be drawn from these districts. For the matching purpose available secondary data sources (such as the Census, DHS, NLSS) will be used and data on the parameters including poverty, health infrastructure, educational measures, utilization of health care, average catchment area of the health facilities, etc. will be utilized for identifying the districts that will be used as *pseudo-control* districts.

[20] Additionally, we might consider using household matching in case there are large variation in outcome measures within a VDC depending on the necessity after the baseline data

Stage 2: MatchingVDCs

In this stage community level surveys will be undertaken to collect VDC level data which will be utilized for VDC matching. From 3 treatment districts, VDC level community survey of 153 VDCs will be undertaken and similarly, VDC level community survey will be undertaken in all the VDCs (approximately 200 VDCs) in the 3 *pseudo-control* districts, leading to a total 350 VDC level community survey to collect data on factors that might affect utilization of health services and out of pocket expenditure.
Data will be collected on the following variables including availability of health infrastructure - public and private, distance and travel time to nearest health facility, to district capital/town, perception about utilization of health facilities by community members. While collecting such data a retrospective series on these variables will be created using recall method which will be utilized for tracking trend in the two groups. Using the community level data, a propensity score matching method will be used to match VDCs from treatment districts with those in the control districts in order to identify pseudo-control VDCs.

Power calculations indicate that it is sufficient to have 100 treatment VDCs to be matched with 100 *pseudo-control* VDCs. So, we can randomly select 100 treatment VDCs from 153 VDCs of the 3 treatment districts and then identify 100 matching *pseudo control* VDCs from a pool of 200 *pseudo control* VDCs in the 3 *pseudo-control* districts.

The impact of the program will be estimated using difference-in-differences with the matches, comparing the before-after program change in each outcome indicator for the treatment VDCs with the same change in the matched, non-treatment VDCs. Given the voluntary nature of the program, this method will produce an estimate of the *intent-to-treat effect*.

Process Evaluation

The process evaluation will include a qualitative component directed at subscribers and non-subscribers of the health insurance. It is comprised of three components: administrative, operational, and financial. The evaluation will be tailored to the precise insurance design characteristics. It potentially covers, but is not limited to, the following dimensions:
Component 1: *Administrative*

- Communication strategy (quality of communication tools for members describing benefits and insurance modalities)
- Member satisfaction (extent to which the scheme measures member satisfaction)
- Membership management and monitoring tools (evaluation of the existence and quality of these tools)
- Use of membership and management tools (extent to which membership and management tools are used to produce management information and how that information is used in management decisions)
- Premium collection monitoring (existence and quality of tools for premium collection and membership fees)

- Verification of benefit entitlement (existence and implementation of procedures carried out by health providers and health insurance to verify entitlement to benefits)
- Claims management information system, claims monitoring and risk portfolio monitoring (quality of claims database, quality of management information produced, and extent of management decisions taken)
- Financial and accounting records (quality of budgeting, cash flow and investment monitoring, quality of accounting records and accounting procedures)
- Membership enrolment (evaluate the risk of adverse selection following the way in which members are enrolled)
- Quality of risk portfolio (determine the exposure to financial risk associated with the type of benefits covered)
- Average claims cost (trends, comparison with non-beneficiaries)
- Claim rejection ratio (number of claims rejected over all claims reported)

Component 2: *Operational*

- Overall membership growth rate (new members in period n over members in period n-1)
- Renewal rate (number of renewals over number of potential renewals)
- Penetration rate (percentage of beneficiaries in target population enrolled according to business plan of health insurance)
- Premium collection rate (share of premiums due collected)
- Average time for provider reimbursement (time elapsed between treatment and payment of provider)

Component 3: *Financial*

- Solvency ratio (admitted assets over liabilities)
- Liquidity ratio (cash over short term payables)
- Net income ratio (net income over earned premium)
- Incurred expense ratio (incurred expenses over earned premium)
- Incurred claims ratio (incurred claims over earned premium)

VI. Data and Outcome Indicators

At least two rounds of household including a baseline and end line data collection, will be carried out to obtain primary data on key outcome indicators. The household data will be complemented by program data and provider level data. All data collection will be contracted out to a qualified firm with a sound experience in conducting household surveys in Nepal. Data quality will be ensured through a series of key quality control measures including:

- the contracting of a data quality specialist
- re-surveying by the data collection firm for an explicit number of households

- computer-assisted field data entry
- real time monitoring of data collected in the field

In addition to the primary data collection at the household level, existing data sources including the Health Management Information System (HMIS) which provides regular monitoring data, service tracking surveys, and household surveys will be used. The recently conducted Nepal living Standard Survey and Nepal Demographic and Health surveys will be used to match the districts and verify the robustness of the evaluation design. Data from HMIS will also be used for regular monitoring of the implementation of the intervention.

The evaluation will also include the collection of cost data, which will be combined with the impact estimates to assess the cost-effectiveness of the intervention and fiscal implications of taking it to scale.
The evaluation will assess the interventions impact along four indicators:

- utilization of inpatient care
- utilization of out-patient care
- Out of Pocket Expense
- Drug stock out

VII. Sampling and Power Analysis

Experimental design

Given the voluntary nature of health insurance enrolment in the treatment districts, we propose using encouragement design to affect the take up of health insurance and then use differential take up for constructing IV estimate (LATE) to assess the impact of health insurance on utilization and out of pocket expenditure.
The encouragement intervention will include all the 153 VDCs of the treatment districts, which will be divided into treatment and control groups for the purpose of encouragement intervention.

Given that with the encouragement intervention, only a certain percentage of the treatment sample will enroll in health insurance, and at the same time, a certain percentage of the control sample will also enroll in health insurance, the encouragement will lead to imperfect compliance of the health insurance take up. So, for the power calculation, we need to address this possibility of imperfect compliance while calculating the sample size requirements.

We expect our encouragement design will be able to induce a health insurance take up differential in the range of 35% (high) - 20% (low).
For the purpose of power calculation we intended to use the out-patient utilization rate of the household as the outcome variable of interest; however, given that we don't have any reference data for this outcome

variable, we utilize data on doctor consultation in the event of acute illness[21] as a proxy for power calculation.

The Nepal Living Standard Survey 2010 data indicates that the doctor consultation rate in the event of acute illness in rural Nepal is approximately 21%. We conduct the power analysis for 3 hypothesized differences (30, 25 and 15 percentage points) in doctor consultation rate between the treatment and control sample since we expect that the provision of health insurance could induce additional utilization rate in the range of 15-30 percentage points.

Given the expected effect size that we would like to detect, we then derived how standardized effect size estimate changes for various rates of take up differential of insurance (the proportion of treatment household who take up insurance minus the proportion of control households who take up insurance among the households in the health insurance treatment districts). We used power calculations assuming a cluster randomized trial [22]to identify the sample size requirements for detecting the various standardized effect sizes that are generated by given values of take up differential of insurance.
Based on few combinations of a given sample size and effect size tradeoff under imperfect compliance, we propose to sample total 6120 households, selecting 40 households from each of the 153 VDCs that will be able to detect the expected range of utilization rate differential for a reasonable range of health insurance take up differential (30% - 15%). The description of power calculations has been presented in the appendix. The following table summarizes the findings from the power calculation.

Table 3: Summary of Power Calculation: Encouragement Design with Imperfect Compliance

Total Number of VDCs=153 (77 treatment, 76 control) Intra-cluster correlation=0.1, 40 household per cluster with total sample size =6120	
Health Insurance take up differential between treatment and control sample (% of treatment household who take up health insurance - % of control households who take up health insurance)	Ability of the sample size to detect any utilization rate difference of
35%	19 percentage points or more
30 – 25%	22-26 percentage points or more
25- 20%	30 percentage points or more

[21] Doctor consultation data is available from the NLSS's recent round (2010) report available at http://microdata.worldbank.org/index.php/catalog/1000/download/20615.

[22] Randomization is done at the VDC level, with intra-VDC correlation of 0.1 and standard assumption of power 80%

We expect that our encouragement design will lead to a health insurance take up differential of at least 25% and thus, the sampling plan (40 households from each of the 153 VDCs) will be able to detect utilization rate difference of 26 percentage points or more.
For the unlikely event of encouragement design failing to generate a take up differential of at least 25% or more, a provision has been made to utilize a quasi-experimental design. Under this approach, a difference- in- difference estimation will be undertaken with a matching of VDCs from the treatment districts with the VDCs from comparable districts that won't receive the health insurance intervention to assess the impact of health insurance on the outcome variable as described in the following section.

Quasi- experimental design

For identifying the sample size requirements for quasi-experimental design, we follow a methodology where the required sample size has been calculated in two stages: in the first stage, the standard cluster randomized trial based power calculation is used to determine the sample size and then in the second stage, the control sample size is significantly inflated (to compensate for the possible loss of control sample while matching) to determine the sample size for standard power consideration. [23]

Based on literature on take up, we expect that the take up of health insurance in the treatment districts will be in the range of 40 – 50% and, we desire to identify utilization differential of 20 percentage points between treatment and pseudo-control districts. We have used the methodology that we presented in the previous section where the sample size calculation includes the possibility of imperfect treatment compliance.
To arrive at the sample size requirements for the quasi-experimental design, we have conducted power calculation with various combinations of treatment and pseudo-control VDCs for various expected health insurance take up rate in treatment districts (please refer to part B in Appendix on Sample Size Calculation).
Table 4 summarizes the power calculations with 100 treatment VDCs matched with 100 *pseudo-control* VDCs (which leads to most economical results to identify the desired treatment effect size of 20 percentage point increase in utilization) as presented below:

Table 4: Power Calculation for Quasi-Experimental Design

Total VDCs=200 with 100 Treatment and 100 *pseudo control* VDCs (Intra cluster correlation =0.1, Power=80%)			
Average Health Insurance Take up in treatment districts	Households per VDC required to detect the utilization difference of 20 percentage points	Total Households to be surveyed from the control VDCs as per outcome of the cluster randomized design	Total Households to be surveyed including the buffer households from the 100 *pseudo control* VDCs
50%	4	400	800
40%	7	700	1400

[23] http://blogs.worldbank.org/impactevaluations/power-calculations-for-propensity-score-matching

As per the Optimal Design output for standard cluster randomized trial, (assuming intra-cluster correlation=0.1 and power =80%) with total 200 VDCs (100 treatment matched with 100 *pseudo-control* VDCs), one needs to sample 4 households and 7 households per cluster for health insurance take up of 50% and 40% respectively for detecting the desired difference in utilization rate (at least 20 percentage point) to achieve a power of 80%. Again, to be conservative, we use the lower range of the expected take up (40%) to arrive at the sample size requirements which requires a total 700 households to be surveyed from 100 pseudo-control VDCs to detect the desired utilization differential. However, as explained in the appendix, incorporating the possibility losing control sample in the matching process, we propose sampling 14 households per control VDC (keeping a sample which is 2 times more than the power calculation based required number of households) to keep sufficient buffer of control households.
Thus, our sampling plan indicates that for the quasi-experimental design, total 1400 households will be surveyed from the 100 matched *pseudo-control* VDCs.
Additionally, matching might also result in power gains due to partial explanation of variation in outcome variables, however, to be conservative, it is not factored in the power calculations presented above. While assessing the impact of health insurance on the important outcome variables using difference-in-difference method, we will utilize various baseline controls to increase the efficiency of the estimator which will increase the power by reducing the residual variance.

VIII. Dissemination Plan and Engagement Strategy

There is strong demand for evidence on the effectiveness and affordability of the health insurance intervention pilot in Nepal. As such, the lessons learned from the process and impact evaluation of the pilot will play an important role in informing the current health insurance policies under development throughout the country.

Since the program will be implemented by the government, relevant government officials will be continuously engaged in the design and implementation stages of the intervention and the impact evaluation. This will ensure the engagement of the policymakers in the impact evaluation and the use of the results for decision making purposes. Early efforts to encourage engagement include participation of the team and select Government of Nepal stakeholders in the SIEF workshop held in Delhi in March.

The team plans to organize additional workshops and engagement activities throughout the course of the project cycle, including a pre-baseline impact evaluation design mission and training; a dissemination activity after the collection of the baseline data and preparation of the baseline report; and a final dissemination activity once the results of the evaluation are available. This final event will focus on the policy relevant aspects of the evaluation to ensure that the results a relevant for the purposes of designing health policies in Nepal.

IX. Research Team

The evaluation team will consist of the following individuals:

Principal Investigator:

Santadarshan Sadhu, PhD, Household Finance Specialist, IFMR Finance Foundation , Chennai, India

Task Team Leaders

- Tekabe Belay, Senior Economist, SASHN, World Bank

Technical and Implementation Support

- Tomas Lievens, Director, Social Policy, Oxford Policy Management

Government of Nepal Primary Counterpart

- Khabiraj Khanal, Ministry of Health and Population, Nepal

Impact Evaluation Coordinator(s)

- Manav Batterai, Health Specialist, SASHN, World Bank

The cost analysis will be supported by Tekabe Belay and a costing specialist to be identified at a later date.

X. Timeline

Activity	Time period
Baseline Surveys	March 2014-May 2014
Encouragement Intervention	June 2014
Implementation of the program	July2014
Process evaluation	June 2014- January 2016
End line survey	February 2016

XI. Ethical review

Nepal Health Research Council (NHRC) is a statutory and autonomous body which provides consent to do study and research related to health. Hence a clearance will be obtained from NHRC before a research commences and this entails a small approval fee. The Institutional Review Board is under NHRC and the research proposal will be submitted as per the NHRC research proposal format. It will be submitted as soon as the funding for the proposal is secured.

Given that the PI is a not a Nepali citizen, we will need institutional clearance from academic and related institution where the PI is affiliated with. We are also required to have at least one Nepali co-investigator, which we already have.

XII. Budget and Funding Justification

SIEF resources will fund one baseline and one follow-up survey to be conducted at the household and facility level. They will also fund the time and travel for the evaluation team and consultants including a costing specialist, a qualitative researcher, and a data quality specialist. The costs for these will be included in the final technical proposal. The Government of Nepal will be funding the health insurance scheme itself.

XIII. Disbursement

Funding for evaluations will be transferred based on the achievement of explicit and pre-established milestones. The following schedule of tranches will apply for all new evaluations. Ongoing evaluations will have a distinct and customized schedule of tranches.

- 55% after regional concept note review and technical review of the IE technical proposal
- 5% after uploading the baseline data and documentation to the Microdata Management Toolkit. (MMT)
- 30% upon submission of the baseline report and report validating the evaluation design
- 10% after uploading the endline data, and endline documentation to the MMT.

Appendix A: Demographics for Five Targeted Districts

Region	District	# of VDCs	Pop.	Total # of HH	Avg # of HH per VDC	Avg HH Size	Literacy Rate (6 yrs and older that read & write)	% of Economic Activity (total population across all households)
Eastern Region	Ilam	49	282,107	54,561	1,113	5.25	60%	74%
Western Region	Baglung	60	268,240	53,555	893	5.08	55%	73%
Far West Region	Kailali	44	614,691	94,395	2,145	6.78	43%	71%
Total		**153**	**1,165,038**	**202,511**				

Appendix B: List of Drugs Included in the Benefit Package

1 Albendazole Tab.
2 Aluminium hydroxide + Magnesium hydroxide Tab.
3 Amoxyciline Tab., Cap.
4 Calamine lotion
5 Chloramphenicol Applicaps
6 Chlorpheniramine Tab.
7 Ciprofloxacin Drops
8 Ciprofloxacin Ointment
9 Clove oil
10 Compound solution of Sodium lactate (Ringers' Lactate) Inj.
11 Ferrous salt + Folic acid Tab.
12 Gamma benzene hexachloride Cream.
13 Gentamycin Inj.
14 Hyoscinebutylbromide Tab.
15 Lignocaine Inj.
16 Magnesium Sulphate Inj.
17 Metoclorpropamide Inj.
18 Metronidazole Tab., Sus.
19 Oral Rehydration Solution (ORS) Powder
20 Oxytocin Inj.
21 Paracetamol Tab., Inj., Syp.
22 Pheniramine Inj.
23 Povidinelodine Solution
24 Sulfamethoxazole + Trimethoprim Tab., Sus.
25 Vitamin B complex Tab.
26 Chloramphenicol Cap., Powder, Sus.
27 Dextrose Solution Inj.
28 Phenobarbitone Tab.

Appendix: Sample Size Calculation

A. Experimental design

Estimating Standardized Effect Size for Encouragement Design with Partial Compliance

Given the voluntary nature of health insurance enrolment in the treatment districts, we propose using encouragement design to affect the take up of health insurance and then use differential take up for constructing IV estimate (LATE) to assess the impact of health insurance on utilization and OOP expenditure. The encouragement intervention will include all the 153 VDCs of the treatment districts, which will be equally divided into treatment and control groups for the purpose of encouragement intervention.

Under the encouragement intervention we can derive **the Standardized Treatment Effect Size (STES)** as a function of take up differential and utilization rate differential as**:**

STES= $[(p_T-p_C)* (xr-x)]/ sqrt(x(1-x)$ ……..……. (1)

where p_T is the proportion of treatment household who take up insurance and p_C is the proportion of control households who take up insurance, xr =utilization rate of household with insurance and x is the utilization rate of household without insurance (baseline value), so (xr-x) is the difference in utilization rate.

For the purpose of power calculation the outcome variable of interest would be the out-patient utilization rate of the household; however, given that we don't have any reference data for this outcome variable, we utilize data on doctor consultation in the event of acute illness[24] as a proxy for power calculation.

The Nepal Living Standard Survey 2010 data indicates that the doctor consultation rate in the event of acute illness in rural Nepal is approximately 21%. We conduct the power analysis for 3 hypothesized differences (30, 25 and 15 percentage points) in doctor consultation rate between the treatment and control sample since we expect that the provision of health insurance could induce additional utilization rate in the range of 15-30 percentage points.

Power Calculation: Encouragement Design

In the first stage of power calculation, we calculate the STES for the case of perfect compliance (where $p_T-p_C=1$) using equation (1) for the hypothesized values of difference in doctor consultation rates.

[24] Doctor consultation data is available from the NLSS's recent round (2010) report available at http://microdata.worldbank.org/index.php/catalog/1000/download/20615.

In the second stage, we estimate how the value of STES changes as a result of imperfect compliance for a given difference in utilization rate. As the treatment compliance is not perfect it will reduce the effect size, so to detect a given hypothesized difference in utilization rate, we need to be able to detect a revised STES that is smaller than the STES with perfect compliance. The following table shows relationship between the STES for perfect compliance and STES that is required to be detected for imperfect compliance for the hypothesized values of difference in utilization rate.

Table A1: Treatment compliance and Implied Treatment Effect Size

				Standardized Treatment Effect Size <u>with Complete compliance</u> =[(xr-x)]/ sqrt(x(1-x))			**Revised Standardized Treatment Effect Size (STES) =[(pT-pC)* (xr-x)]/ sqrt(x(1-x)) that needs to be detected for partial compliance**			
Baseline utilization (x)	Difference in utilization rate between Insured and uninsured households x(r-1)			x(r-1) =0.3	x(r-1) =0.25	x(r-1) =0.15	Take up differential (Pt-pc)	x(r-1) =0.3	x(r-1) =0.25	x(r-1) =0.15
0.21	0.3	0.25	0.15	0.737	0.614	0.368	**0.5**	0.368	0.307	0.184
0.21	0.3	0.25	0.15	0.737	0.614	0.368	**0.4**	0.295	0.246	0.147
0.21	0.3	0.25	0.15	0.737	0.614	0.368	**0.35**	0.258	0.215	0.129
0.21	0.3	0.25	0.15	0.737	0.614	0.368	**0.3**	0.221	0.184	0.110
0.21	0.3	0.25	0.15	0.737	0.614	0.368	**0.25**	0.184	0.153	0.092
0.21	0.3	0.25	0.15	0.737	0.614	0.368	**0.20**	0.147	0.123	0.074

As can be seen from the above table as the compliance rate falls, the STES required to detect a given utilization rate differential also falls, indicating that the sampling plan should be robust enough to be able to detect smaller and smaller effect size.
Finally, we run the power calculations (using standard OD program) to explore the STES that the given sample size can detect from 153 VDCs (clusters, with a given power of 80% and intra-cluster correlation of 0.10). Using the OD program, we explore the range of STES that can be detected from various sample sizes. Based on that exercise, we propose to sample 40 households per VDC leading to a total sample size of 6120 which will be able to detect STES of 0.16.

Based on the template of the above table we can summarize the effectiveness of our suggested sampling plan as follows:

(i) For any encouragement design that results in a health insurance take up differential of at least 35 percentage points between the treatment and control sample , the sampling plan will be able to detect utilization rate difference of 19 percentage points or more

(ii) For any encouragement design that results in a health insurance take up differential of 30 percentage points between the treatment and control sample , the sampling plan will be able to detect utilization rate difference of 22 percentage points or more

(iii) For any encouragement design that results in a health insurance take up differential at the range of 25 percentage points between the treatment and control sample , the sampling plan will be able to detect utilization rate difference of 26 percentage points or more

(iv) For any encouragement design that results in a health insurance take up differential of more than 20 percentage points (and less than 25 percentage points) between the treatment and control sample , the sampling plan will be able to detect utilization rate difference of 30 percentage points or more.

Table A2: Effectiveness of Suggested Sample Size

	Total Number of VDCs (3 districts) =153 (77 treatment, 76 control) **Intra-cluster correlation=0.1, 40 household per VDC** & **Sample Size =6120**
Treatment compliance ($p_T - p_C$) in percentage points	Ability of the sample size to detect any utilization rate difference of
0.35	19 percentage points or more
0.3	22 percentage points or more
0.25	26 percentage points or more
More than 20 –less than 25	33 percentage points or more

We expect that our encouragement design will lead to a take up differential of at least 25% and thus, the **sampling plan** (40 households from each of the 153 VDCs) will be able to detect **utilization rate difference of 26 percentage points.**

For the unlikely event of encouragement design failing to generate a take up differential of at least 25% or more, a provision will be made to utilize a quasi-experimental design where a difference- in- difference estimation will be undertaken to assess the impact of health insurance on the outcome variable as described in the following section.

B. Quasi- experimental design

Under this approach, we propose using a matching of VDCs from comparable districts that won't receive the health insurance intervention (*pseudo-control districts*) with the VDCs from the treatment districts.

Power Calculation: Quasi-experimental design with Matching

Since there exists no standard and universally accepted measure for conducting the power calculation, for the quasi-experimental design, we utilize a method suggested by David McKenzie in one of his blogs.[25] Using the suggestion from the blog, we follow a methodology where the sample size for a quasi-experimental design can be calculated in two stages: in the first stage, the standard cluster randomized trial based power calculation is used to determine the sample size and then in the second stage, the control sample size is significantly inflated (to compensate for the possible loss of control sample while matching) to determine the sample size for standard power consideration.

[25] http://blogs.worldbank.org/impactevaluations/power-calculations-for-propensity-score-matching

For our case with the encouragement design, while attempting to use the cluster randomized method for calculating the sample size, we first recognize that there is a partial compliance of treatment in the treatment districts (take up of health insurance in the treatment districts will be less than 100%), indicating that we need to use the insights from the previous section to derive the sample size incorporating the imperfect compliance possibility.
Based on literature on take up, we assume that on average, the take up of insurance in the treatment districts will be in the range of 40 – 50% and we expect utilization differential of 20 percentage points between treatment and pseudo-control districts.

To arrive at the sample size requirements for the quasi-experimental design, we have conducted power calculation under different scenarios with differing numbers of treatment and *pseudo-control* VDCs (with maximum 153 treatment to be matched with 153 *pseudo-control* VDCs and a minimum of 100 VDC from treatment districts (selected randomly from the 153 VDCs) to be matched with 100 VDCs from the control districts) and presented the calculation for the case with 100 treatment matched with 100 *pseudo-control* VDCs as it leads to most economical results to identify the desired treatment effect size (20 percentage point increase in utilization) as presented below:

Table A3: Power Calculation for Sample Size Determination for Quasi-Experimental Design

Total VDCs=200 with100 Treatment, 100 Matching control VDCs (Intra cluster correlation =0.1,Power =80%)			
Average Health Insurance Take up in treatment districts	**Revised Standardized Treatment Effect Size (STES) =[(pT-pC)* (xr-x)]/ sqrt(x(1-x)) that needs to be detected for partial compliance**	Households per VDC required to detect the utilization difference of 20 percentage points	Total Households to be surveyed including the buffer from the control VDCs
50%	0.246	4	800
40%	0.196	7	1400

As per the Optimal Design output for standard cluster randomized trial, (assuming intra-cluster correlation=0.1 and power =80%) with total 200 VDCs (100 treatment matched with 100 *pseudo-control* VDCs), one needs to sample 4 households and 7 households per cluster for health take up differential of 50% and 40% respectively for detecting the desired difference in utilization rate (20 percentage point) to achieve a power of 80%. Again, to be conservative, we use the lower range of the expected take up (40%) to arrive at the sample size requirements.
However, we suggest keeping a provision of using household matching in case we find a large inter VDC variation in outcome variables. While undertaking a household match, there is a possibility that we might lose control sample (those beyond the common support range of matching), and thus to incorporate that possibility we propose keeping sufficient buffer in control households. Following the suggestion from David McKenzie's blog, we propose sampling 14 households per control VDC (keeping a sample which

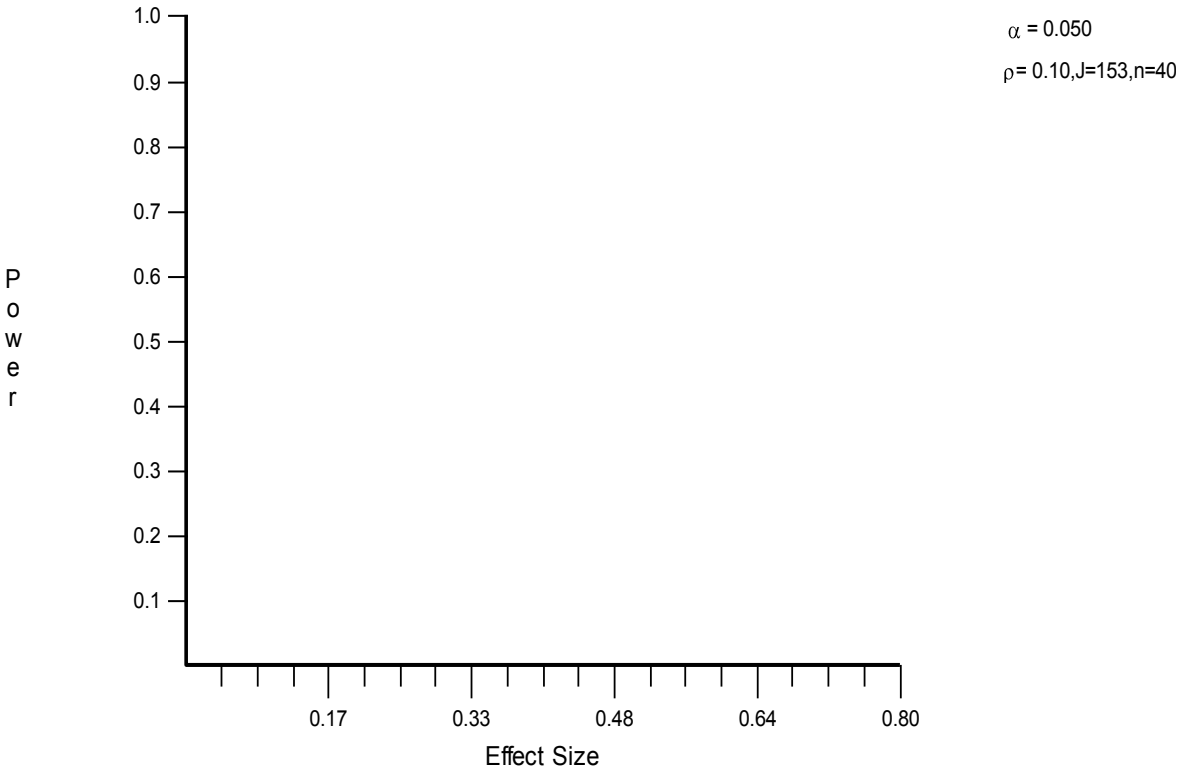
is 2 times more than the power calculation based required number of households leading to sufficient buffer of households for match). This indicates that for the quasi-experimental design, total 1400 households will be surveyed from the 100 matched *pseudo-control* VDCs. Additionally, matching might also result in power gains due to partial explanation of variation in outcomes, however, it is not factored in the power calculations presented above.

While assessing the impact of health insurance on the important outcome variables using difference-in-difference method, we will utilize various baseline controls to increase the efficiency of the estimator which will increase the power by reducing the residual variance.

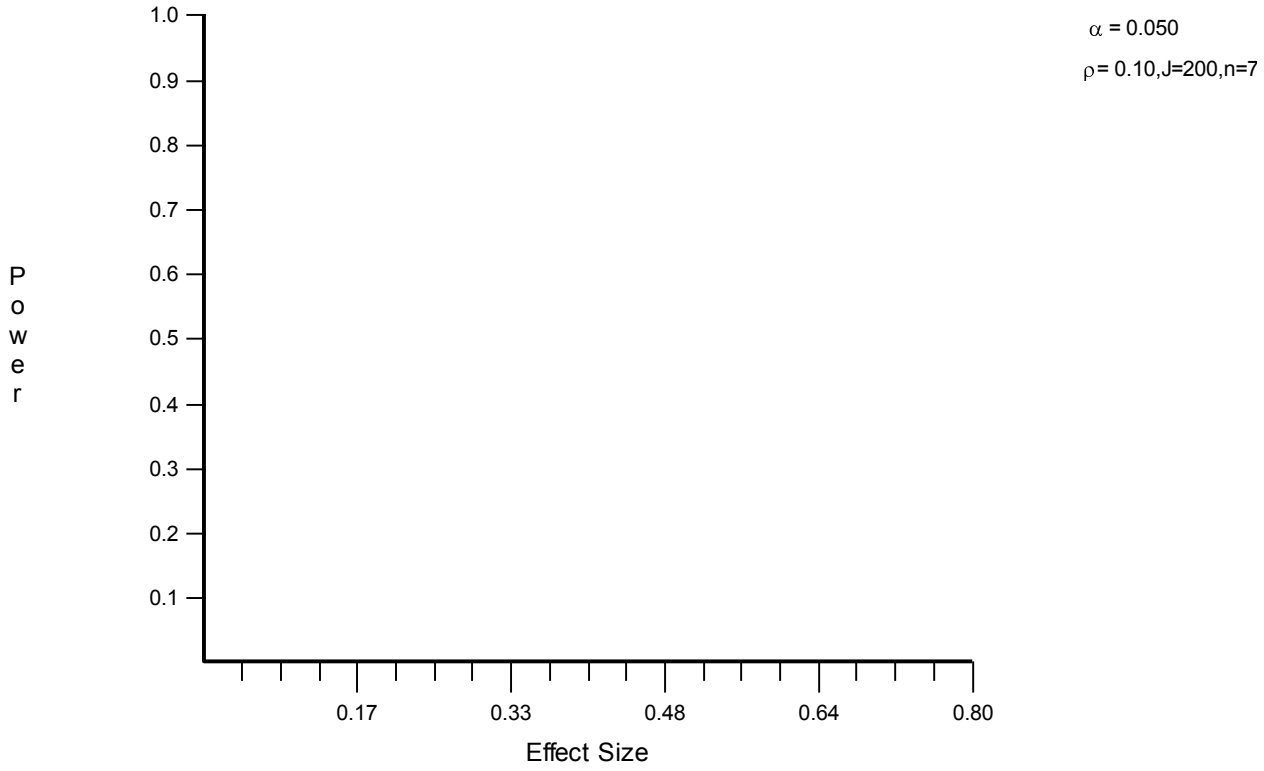
Appendix D:Power Calculation: Using Optimal Design

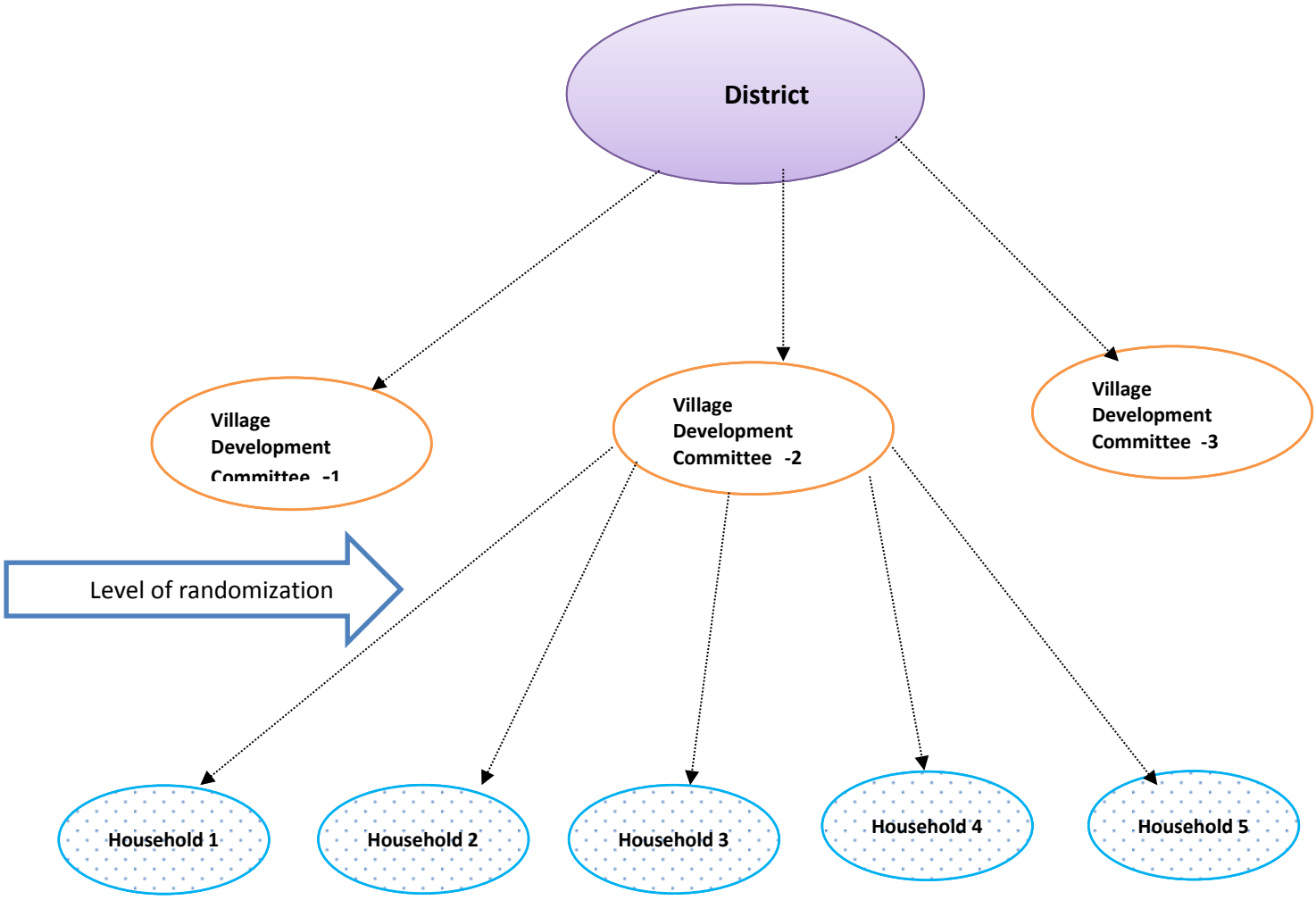
Experimental Design

Cluster Randomized Trail with 153 Clusters (VDCs)
Effect Size vs. Sample Size Per Cluster



Quasi-Experimental Design
Cluster Randomized Trail with 200 Clusters (VDCs)
Effect Size vs. Sample Size Per Cluster




District
Village Development Committee -1
Village Development Committee -2
Village Development Committee -3
Level of randomization
Household 1
Household 2
Household 3
Household 4
Household 5

Annex I: Requirements

All evaluation teams must agree to the following requirements:

1. Ethics Principles

SIEF evaluation teams must adhere to the highest standards of research ethics and are thereby expected to uphold the Ethics Principals for Research and Evaluation established by the United Kingdom's Department for International Development and adopted by the SIEF.

ETHICS PRINCIPLES FOR RESEARCH AND EVALUATION

1. **Researchers and evaluators are responsible for identifying the need for and securing any necessary ethics approval for the study they are undertaking.** This may be from national or local ethics committees in countries in which the study will be undertaken, or other stakeholder institutions with formal ethics approval systems.

2. **Research and evaluation must be relevant and high quality with clear developmental and practical value.** It must be undertaken to a sufficiently high standard that the findings can be reliably used for their intended purpose. Research should only be undertaken where there is a clear gap in knowledge. Evaluations might also be undertaken to learn lessons to improve future impact, or in order to meet DFID's requirements for accountability.

3. **Researchers and evaluators should avoid harm to participants in studies.** They should ensure that the basic human rights of individuals and groups with whom they interact are protected. This is particularly important with regard to vulnerable people. The wellbeing of researchers/ evaluators working in the field should also be considered and harm minimized.

4. **Participation in research and evaluation should be voluntary and free from external pressure.** Information should not be withheld from prospective participants that might affect their willingness to participate. All participants should have a right to withdraw from research/ evaluation and withdraw any data concerning them at any point without fear of penalty.

5. **Researchers and evaluators should ensure confidentiality of information, privacy and anonymity of study participants.** They should communicate clearly to prospective participants any limits to confidentiality. In cases where unexpected evidence of serious wrong-doing is uncovered (e.g. corruption or abuse) there may be a need to consider whether the normal commitment to confidentiality might be outweighed by the ethical need to prevent harm to vulnerable people. DFID's fraud policy will apply if relevant.

6. **Researchers and evaluators should operate in accordance with international human rights conventions and covenants to which the United Kingdom is a signatory, regardless of local country standards.** They should also take account of local and national laws.

7. **DFID funded research and evaluation should respect cultural sensitivities.** This means researchers need to take account of differences in culture, local behavior and norms, religious beliefs and practices, sexual orientation, gender roles, disability, age and ethnicity and other social differences such as class when planning studies and communicating findings. DFID should avoid imposing a burden of over-researching particular groups.

8. **DFID is committed to publication and communication of all evaluations and research studies.** Full methodological details and information on who has undertaken a study should be given and messages transmitted should fully and fairly reflect the findings. Where possible, and respecting confidentiality requirements, primary data should be made public to allow secondary analyses.

9. **Research and evaluation should usually be independent of those implementing an intervention or program under study.** Independence is very important for research and evaluation; in fact evaluations in DFID can only be classified as such

where they are led independently. Involvement of stakeholders may be desirable so long as the objectivity of a study is not compromised and DFID is transparent about the roles played. Any potential conflicts of interest that might jeopardize the integrity of the methodology or the outputs of research/ evaluation should be disclosed. If researchers/ evaluators or other stakeholders feel that undue pressure is being put on them by DFID officials, such that their independence has been breached, this should be reported to the Head of Profession for Evaluation who will take appropriate action

10. **All DFID funded research/ evaluation should have particular emphasis on ensuring participation from women and socially excluded groups.** Consideration should be given to how barriers to participation can be removed.

2. Protection of Human Subjects

SIEF evaluations must fulfill the following requirements on human subjects:

- Technical proposals must include a description of the human subjects protocol and a plan for securing ethical clearance in the technical proposal.
- Principal investigators and research coordinators must provide evidence of human subjects training within the last 2 years. Technical proposal should include a list of the ethical research training taken by the PI and co-PI. The National Institute of Health (NIH) online course includes a test and will produce a certificate number which can be used for this purpose.
- Principal investigators are responsible for securing in-country ethical clearance or providing an official memo from client counterparts stating the absence of a local ethical review board. In case the country of study does not have a review board, the evaluation team will be required to contract an external review board.

3. Data Storage and Access

In accordance to the World Bank's Open Data and Open Knowledge Initiative, all datasets must be fully documented. Datasets should be in compliance with international good practices and with the Data Documentation Initiative (www.ddialliance.org).
To promote broad and diverse use of the data, and to ensure transparency and credibility of the results, microdata will be made **publicly accessible** within **two years** from the submission of the final evaluation report. Data and documentation will be stored and documented in the World Bank Data Catalog within six months of completion of data collection.

4. Cost Data and Analysis

SIEF evaluations must include the collection of cost data of the intervention. In addition, the final results coming from the evaluations should include an analysis of both the impacts and the costs.

5. Engagement and Dissemination

SIEF evaluations must outline a dissemination and engagement strategy to ensure the relevance of the impact evaluation, including the preparation of a technical report, a policy note, and a final dissemination event. Research supported by the SIEF is expected to be disseminated through World Bank and other channels.