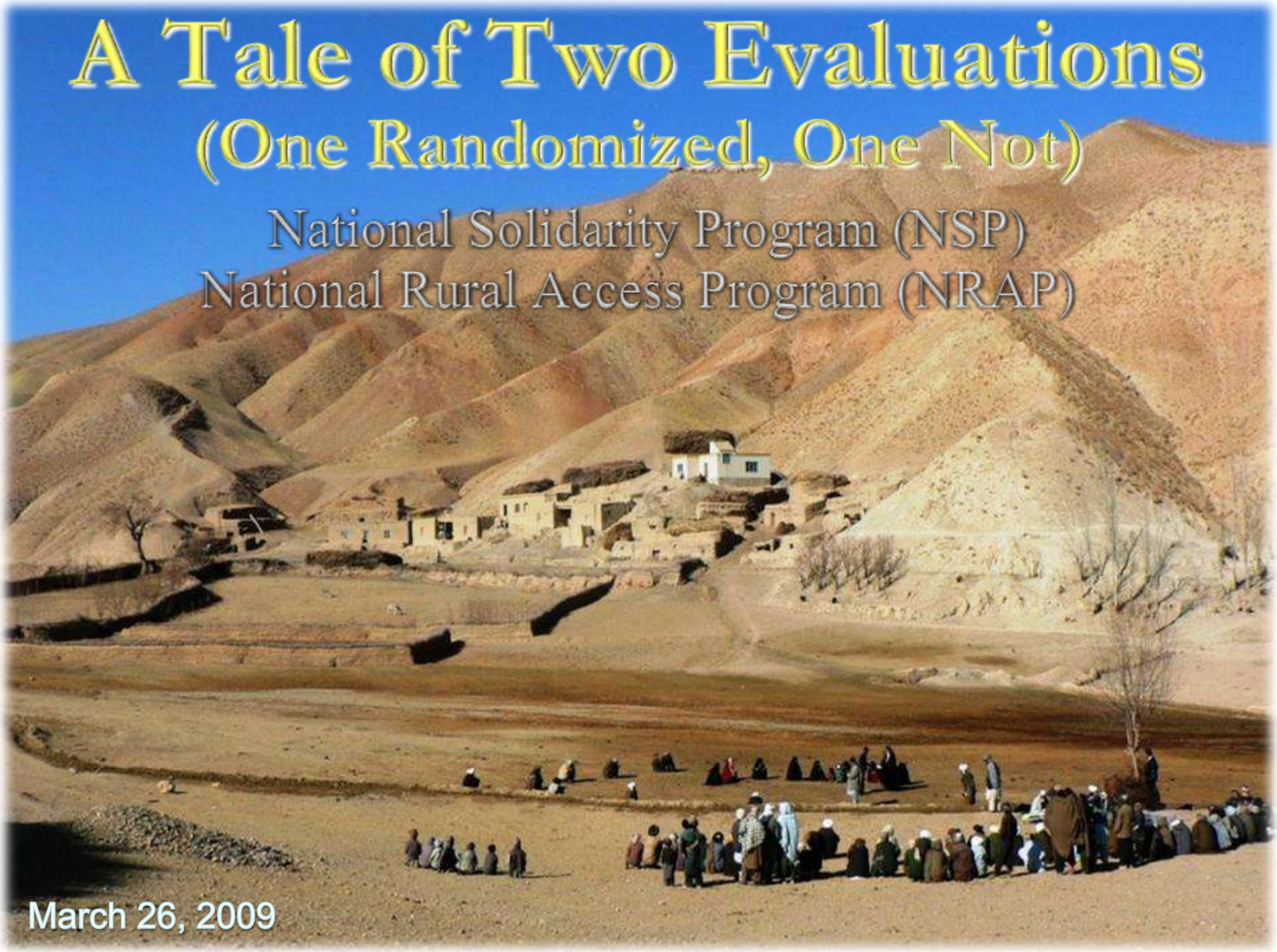
# A Tale of Two Evaluations

## (One Randomized, One Not)

National Solidarity Program (NSP)
National Rural Access Program (NRAP)



March 26, 2009

## Evaluation Teams:

Andrew Beath (Ph.D. Candidate, Harvard University) – NSP-IE & NRAP-IE
Veronica Minaya (World Bank) – NRAP-IE
Shahim Kabuli (World Bank) – NSP-IE & NRAP-IE
Prof. Fotini Christia (M.I.T.) – NSP-IE
Prof. Ruben Enikolopov (New Economic School, Moscow) – NSP-IE
Justin Grimmer (Ph.D. Candidate, Harvard University) – NRAP-IE
Sandra Sequeira (Ph.D. Candidate, Harvard University) – NRAP-IE

## Evaluation Websites:

**NSP-IE:** http://web.mit.edu/cfotini/www/NSP-IE
**NRAP-IE:** http://www.beath.org/NERAP-IE

# National Solidarity Program (NSP)

- Create CDC through secret-ballot election
  - Size of CDC proportional to size of community; ½ male, ½ female
- Build Capacity of CDC Members
  - Community Development Plan (CDP) is drafted; sub-projects are proposed to NSP for funding
- Disburse Block Grants to Fund Sub-Projects
  - $200 per household (community max: $60,000)
- Expand Role of CDCs
  - Form linkages with other agencies and development programs; Governance role of CDCs is yet to be decided

## NSP-I: 2003 – 2007

– 17,300 communities (279 / 398 districts)

## NSP-II: 2007+

– 17,540 communities, split among 74 / 398 "new districts" and 69 "on-going districts"
– In "new districts", NSP is to be rationed to 40 communities / district

# National Rural Access Program (NRAP)

- "Post-Bonn" goal to provide road connectivity to 40% of 38,000 rural villages (19 million people) by end-2010
  - National Emergency Rural Access Program (NERAP), funded by IDA, is component of the larger NRAP program proposing to construct or rehabilitate 4,000 km of rural roads between through 2010
- NERAP aims to "*provide year-round access to basic services and facilities in the rural areas of Afghanistan covered by the project*"
  - KPIs include reductions in travel times, increase in frequency of trips, and reductions in prices of basic commodities
- NERAP constructs or rehabilitates 2,010 km of rural roads and bridges:
  - Improvement of 1,081 km of secondary roads by Ministry of Public Works (MPW)
  - Improvement of 939 km of tertiary roads by Ministry of Rural Rehabilitation and Development (MRRD)
- Selection based on political and economic criteria
  - Selected projects include direct requests from rural communities and parliamentarians, roads considered to increase connectivity to previously isolated areas, and consolidation of previous investments

# Goals of Evaluations

- Provide a rigorous and detailed assessment of program impacts on outcomes of interest (*does the program work when implemented correctly?*);
- Provide evidence to inform decisions made by Government of Afghanistan, donors, and NGOs concerning structure of development assistance;
- Create knowledge which can inform the design of other development programs, both in Afghanistan and other developing countries.

# Functions of Evaluations

**Are villages and households that receive the program in a better situation than they would have been had they not received the program?**

**NSP-IE:**

- How does NSP affect local institutional structures?
- Does NSP build trust among villagers?
- Does NSP increase female participation in governance?
- Does NSP increase access to services?
- Does NSP alter levels and diversity of village consumption and production?

**NRAP-IE:**

- Does NRAP lower price levels of basic commodities?
- Does NRAP diversify sources of food and improve nutrition?
- Does NRAP increase revenues from trade and production?
- Does NRAP increase access to services?
- By how much does NRAP reduce travel times between villages and district centers? By how much does it increase number of trips made by villagers to district centers?
- Do villages that are wealthier or close to district centers benefit more from NRAP projects than villagers that are poorer or further away?

# NSP-IE Methodology

# Feasibility of Randomization

Randomization of NSP was feasible in some districts:

- Resource constraints necessitated rationing of NSP-II to 40 communities across 74 "new districts";
- Absence of village-level economic data limited possibilities to accurately and transparently target program within districts;
- Program and implementing partners were supportive of randomization.

Number of randomization-feasible districts was limited:

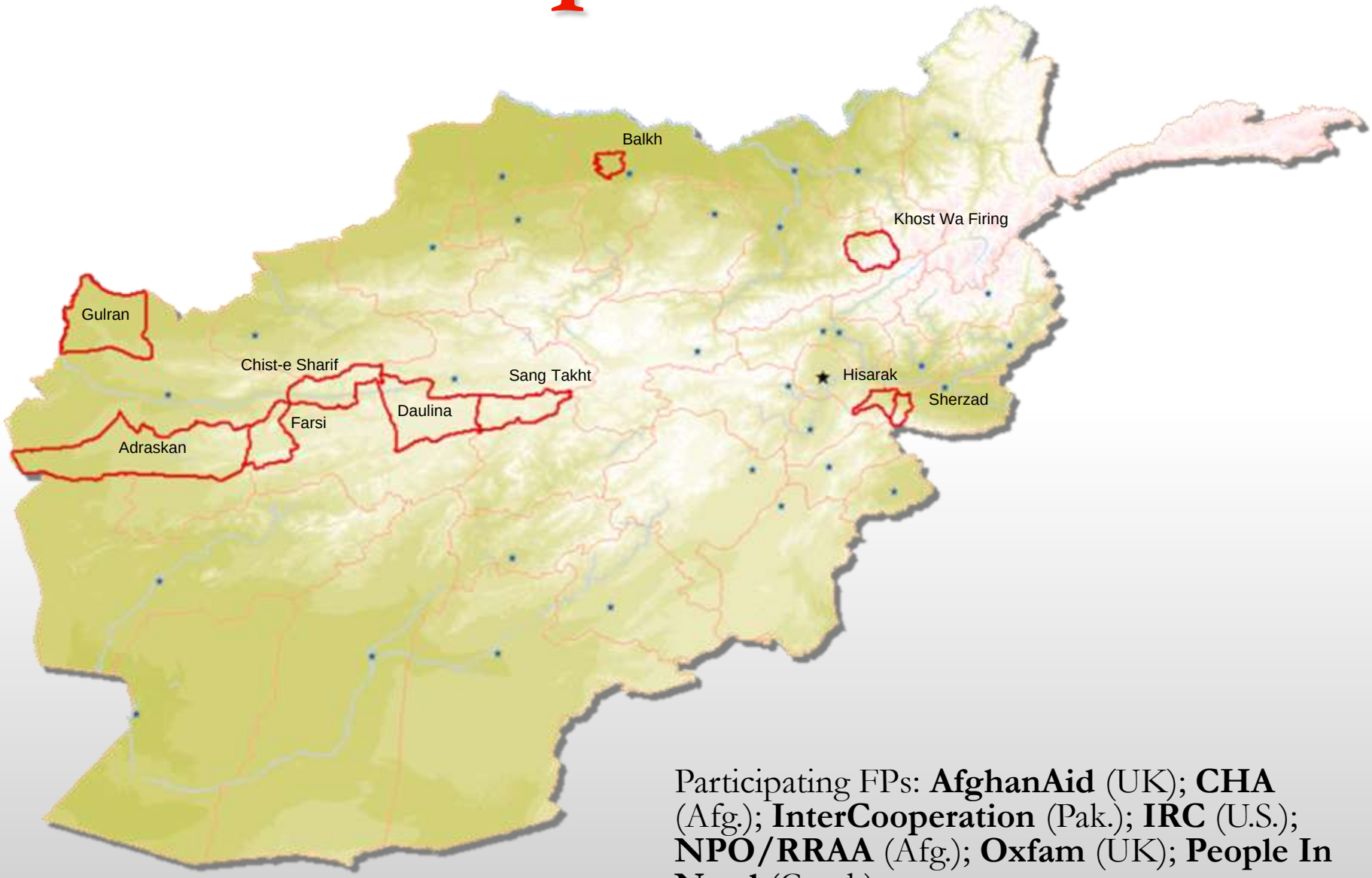
- Districts must have minimum of 65 villages, which eliminated 51 of 74 new districts;
- Districts must be safe for survey activities, which eliminated 12 of the 23 suitably large districts;

10 sample districts are considered to be reasonably representative:

- Ethnic, linguistic, economic, and geographic mix is generally representative of Afghanistan, although for security reasons, districts in southern Afghanistan could not be included;

. . . but limitation of area over which randomization occurred potentially constrains external validity.

# 10 Sample Districts




Participating FPs: **AfghanAid** (UK); **CHA** (Afg.); **InterCooperation** (Pak.); **IRC** (U.S.); **NPO/RRAA** (Afg.); **Oxfam** (UK); **People In Need** (Czech)

# Selection of Sample Villages

1. FPs issued list of villages in district;
2. FPs select 50 sample villages in each district:
   - 50 evaluation villages to be included in evaluation, of which 25 are to be randomly selected to receive NSP, with the remainder forming the control group;
   - Procedure is necessary to ensure evaluation did not create logistical difficulties for participating NGOs
3. FPs select 15 priority villages to receive NSP, but to be excluded from evaluation:
   - Necessary to ensured political and humanitarian imperatives for targeting could be met within context of evaluation;
   - Evaluation team vets lists of 15 priority villages to ensure no overlap with 50 evaluation villages.

# Details of Randomization

1. 25 matched-pairs of villages formed in each of 10 sample districts using multivariate matching:
   - *Attrition Proof*: Protects inferences from loss of sample units;
   - *Interaction Effects*: Permits inferences over interaction of program effects with underlying conditions.
2. Treated assigned to one unit in matched-pair using random number generator:
   - *No Selection Bias*: Negligible differences between units selected to receive program and those selected not to receive program;
   - *Facilitates Transparent Estimation*: Program effects are the difference between outcomes in 250 NSP evaluation villages and the 250 non-NSP evaluation villages.

# Success of Randomization

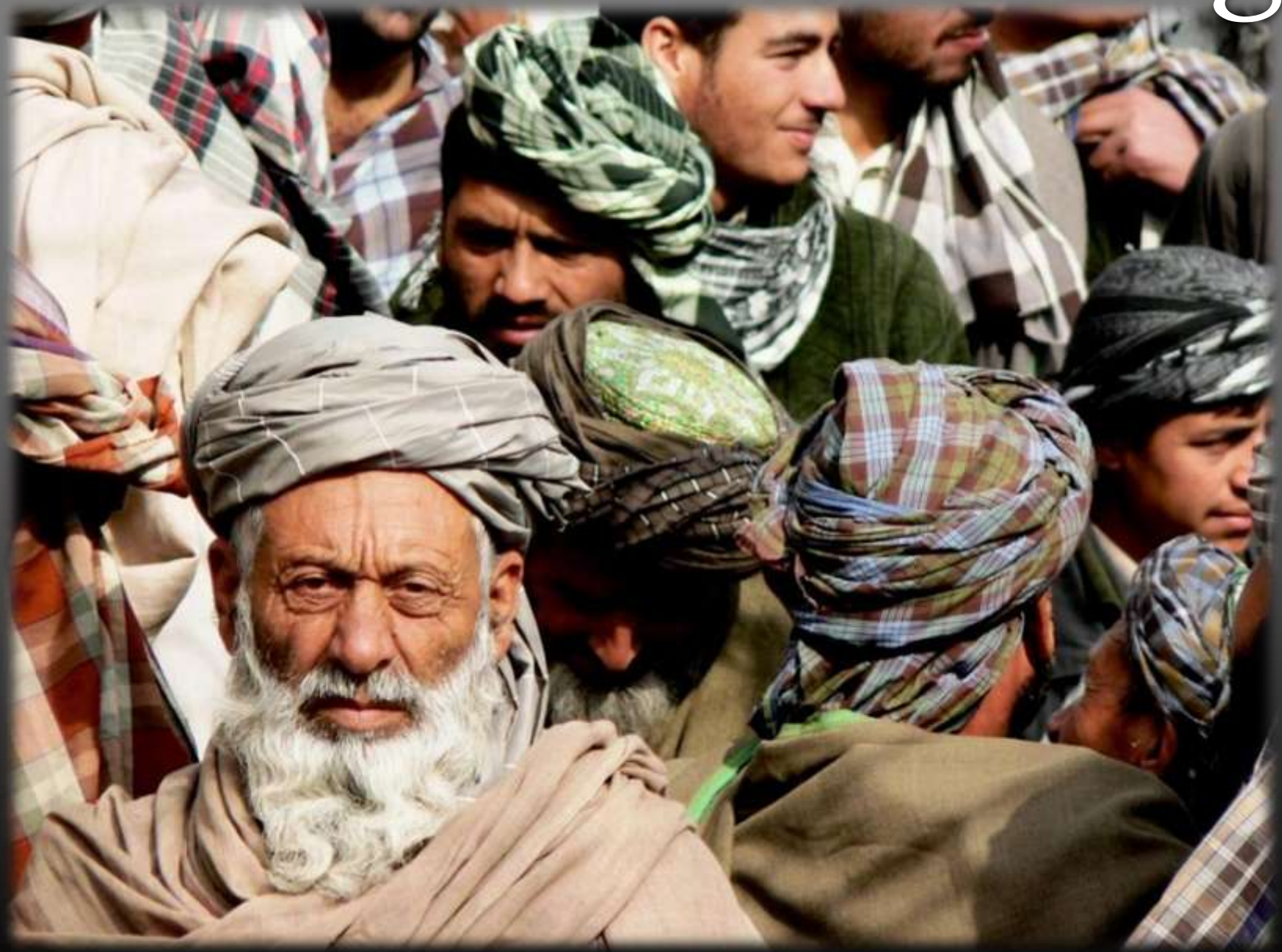
| Variable | Mean for Treatment Group | Mean for Control Group | Normalized Difference |
|---|---|---|---|
| Number of Households | 136 | 148 | **-0.030** |
| Dari Speakers | 67.8% | 66.5% | **0.027** |
| No Education | 69.0% | 68.8% | **0.005** |
| Access to Electricity | 14.1% | 13.0% | **0.033** |
| Dispute in Village | 34.3% | 35.9% | **-0.034** |
| Borrowed Money Last Year | 46.1% | 47.5% | **-0.029** |
| No Access to Medical Services | 12.0% | 12.0% | **0.001** |
| Attended Meeting of Village Services | 30.9% | 29.5% | **0.030** |
| Expenditures on Weddings | $369 | $343 | **0.007** |
| Women Should be Members of Shura | 42.8% | 40.3% | **0.051** |

# Sub-Treatment Interventions (STIs)

Designed to test potential refinements in NSP:

- **Election Type**:
  - 125 / 250 villages elect CDC with ward election
    - Ensures representation for all wards in village
  - 125 / 250 villages elect CDC with at-large election
    - Allows village to elect desired candidates
- **Sub-Project Selection Procedure**:
  - 125 / 250 select sub-projects by village meeting
    - Allows for discussion and debate of best alternatives
  - 125 / 250 select sub-projects by secret-ballot referendum
    - Directly democratic method of selection

# NRAP-IE Methodology



# Alternative to Randomization

- Randomization of NRAP was not feasible as roads had already been selected prior to the design of the evaluation;
- Villages close to selected roads are likely to differ in important characteristics (incl. *ex-ante* outcome indicators) from those villages close to unselected roads:
  - A simple comparison of outcomes or even change-of-outcomes between villages would be contaminated by selection biases etc.
- To compose a suitable 'control group', a statistical matching procedure was designed to identify unselected roads and villages which are similar in important characteristics to selected roads and villages:
  - The procedure, known as propensity score matching, seeks to quantify the selection process by statistically identifying characteristics that increase the propensity of a unit (i.e. a road or a village) be selected to receive the program;
  - A control group is then selected from unselected units which have 'propensity scores' similar to the selected units (these are units which appear to meet the criteria for receiving the program, but were not selected);

# Propensity Score Matching (PSM)

- PSM works best when amount of data on selected and unselected units is large, diverse, of high quality, and includes factors incorporated into selection procedures;
- Afghanistan's sub-national data is limited:
  - Only available source of village-level data available is 'household listing exercise' conducted by Central Statistics Office (CSO) over 2003 – 5, which gives demographic and infrastructure figures for 38,000 villages;
  - Data on characteristics purportedly important for NRAP selection were unavailable and unable to be derived;
  - Fortunately, GPS coordinates of 81 NERAP roads and 'population' of 1,940 primary, secondary, and tertiary roads in Afghanistan exist and, in conjunction with CSO village-level data, were used to generate characteristics for road segments in Afghanistan, from which propensity scores were calculated.
- To best approximate the NRAP selection process, PSM was conducted at the road-level (which incorporated data from surrounding villages) as roads, not villages, are what is selected to receive the program;
  - Separate PSM models were built for secondary and tertiary roads to reflect the different selection procedures;
  - Variables included in PSM were village size, average household size, demographic structure, language, topography, NSP, distance to district center etc.

# Evaluation Structure

- Due to the limited quality of the data used for PSM and selection of the control group, a 2-to-1 matching strategy was employed to select 2 control roads for every treatment road (340 control villages vs. 226 treatment villages):
  - This increased the number of villages to be surveyed by the baseline survey by 50 percent, but enables the evaluation team to utilize the wide array of data provided by the baseline survey to conduct a second-round of PSM;
  - The control group will thereby be reduced by 50% to include those control roads that best 'match' the treatment roads;
  - Discarded control roads will not be included in the follow-up surveys.
- Program effects will be calculated using difference-in-difference estimates between the control and treatment groups using data from the baseline and follow-up surveys.

# Success of PSM

| Variable | Mean for Treatment Group | Mean for Control Group | Normalized Difference |
|---|---|---|---|
| Dari Speakers | 60.8% | 74.4% | **-0.294** (0.027) |
| No Education | 65.3% | 59.0% | **0.123** (0.005) |
| Expenditures on Weddings | $179 | $183 | **-0.004** (0.007) |
| Borrowed Money Last Year | 68.2% | 75.7% | **-0.104** (-0.029) |
| Received Treatment for Last Injury | 82.3% | 85.4% | **-0.036** |
| Heard of NSP | 95.4% | 91.7% | **0.147** |
| NSP Council in Village | 57.6% | 65.4% | **-0.124** |
| Project in Village in Past 5 Years | 89.3% | 87.1% | **0.066** |

# Progress & Timelines

# NSP-IE

**Baseline Survey completed; Program implementation advanced; 1st Follow-Up Survey in May – August 2009:**

- Partnership with MRRD's Vulnerability Analysis Unit (VAU) to collect data for baseline and follow-up surveys;
- Baseline survey of 13,000 people in 500 villages completed in Sept. 2007: Baseline Survey Report available on NSP-IE website;
- Election and sub-project selection monitoring completed in July 2008 (1,500 people in 125 villages interviewed for each): Monitoring Reports available on website;
- Draft STI reports available shortly;
- 1st Follow-Up Survey planned for May – Aug. 2009: Report on interim program impact expected by end-2009;
- Second follow-up survey tentatively planned for Aug. – Sep. 2010.

# NERAP-IE

**Baseline Survey completed; Program implementation starting; 1st Follow-Up Survey in 2010:**

- Partnership with MRRD's Vulnerability Analysis Unit (VAU) to collect data for baseline and follow-up surveys;
- Baseline survey of 13,000 householders (5,660 men and 5,640 women) in 566 villages completed in Oct. 2008: Baseline Survey Report available shortly
- Presentation on NERAP-IE Baseline Survey on March 31 at 12:30pm in MC9-100
- 1st Follow-Up Survey planned for fall 2010.

# General Guidelines

**Costs and timelines vary greatly depending on the program under evaluation and country context;**

**Results will hopefully inform program funding and design decisions made by government and donors, but this is dependent upon buy-in from stakeholders**