USAID | PAKISTAN
FROM THE AMERICAN PEOPLE

Endline Evaluation
Final Report

Pakistan Safe Drinking Water and Hygiene Promotion Project
(PSDW - HPP)

January 2011

This publication was produced for review by the United States Agency for International Development. It was prepared by Abt Associates Inc.

Pakistan Safe Drinking Water and Hygiene Promotion Project

Endline Evaluation

Contract No. GHA-I-00-04-00001-00
Order No. 1

Prepared for USAID
January 2011

Prepared by

Abt Associates Inc.

DISCLAIMER: The author's views expressed in this publication do not necessarily reflect the views of the United States Agency for International Development or the United States Government.

Table of Contents

List of Tables

List of Figures

List of Acronyms

ADB	Asian Development Bank
AED	Academy for Educational Development
AJK	Azad Jammu and Kashmir
DID	Difference-in-Difference Estimate
FATA	Federally Administered Tribal Areas
GOP	Government of Pakistan
NSP	National Sanitation Policy
KPK	North-West Frontier Province
PSDW-HPP	Pakistan Safe Drinking Water and Hygiene Promotion Project
SEC	Socio-economic class
UNDP	United Nations Development Program
UNICEF	United Nations Children's Fund
USAID	United States Agency for International Development
WHO	World Health Organization
WSSD	World Summit on Sustainable Development

1. Introduction

According to the World Health Organization (WHO) and the United Nations Children's Fund (UNICEF), approximately four billion cases of diarrhea are reported worldwide each year, resulting in 1.8 million deaths, or roughly 5,000 deaths per day. This mortality rate is responsible for 15 to 18 percent of all deaths of children under age five, making diarrhea one of the biggest killers of children (WHO, 2007). Unsafe drinking water, inadequate sanitation, and poor hygiene are the main causes of diarrhea in children. Unsafe drinking water also has a disproportionate affect on the poor (ADB 2004; UNDP 2006a). The combination of unsafe water consumption and poor hygiene practices causes hardships for families as it leads to high-cost treatments for waterborne illnesses, causes working days to be lost, and contributes to decreased educational achievement (due to reduced school attendance by children).

In Pakistan, the mortality rate for children under age five is 101 deaths per 1,000 children. Water- and sanitation-related diseases are also responsible for 11% child mortality cases. This rate is expected to have increased since the recent flooding (UNICEF 2006). The underlying causes of diarrheal diseases include inadequate access to safe water, poor household and environmental sanitation, and poor hygiene practices.

Access to safe drinking water is also a critical health issue in Pakistan. The projected population growth for the next 10 years—from 150 million to 221 million people—will only exacerbate water demand, making access to safe drinking water even more of a challenge. Data indicate that 45 percent of the population has access to improved sanitation facilities and 90 percent to improved drinking water sources (Demographic Health Survey, 2006-2007). Delivery of potable water supply is constrained by the inability of local governments, which are now responsible for providing drinking water according to Pakistan's decentralization policy, to manage sustainable water systems (Ahmad, et. al., 2000).

Poor hygiene practices, such as lack of hand washing with soap at multiple critical times, are common in Pakistan. In addition, there is a lack of awareness about what "clean" water means; most people believe that if water is clear and odorless, it is suitable for drinking. This misconception could present a barrier for the acceptance of household water treatment methods or community water filtration plants. Until recent years, environmental health programs have not given behavior change its due importance. A critical mass of the population (more than 66 percent) needs to adopt good water, sanitation, and hygiene behaviors to ensure that public health impacts will, in due course, appear in the district, national, and international statistics (Esrey, 1999).

Pakistan, as a signatory to the Millennium Development Goals, has committed to meeting the targets set at the World Summit on Sustainable Development (WSSD). These targets include halving the proportion of people without sustainable access to safe drinking water and basic sanitation by 2015, which means increasing country-wide drinking water coverage to 93 percent and sanitation coverage to 68 percent by 2015. The Government of Pakistan (GOP) has made allocations in its Medium Term Development Framework (2005–2010) to achieve these targets. Programs supported by various development partners are also being launched to complement the GOP's initiatives. Safe drinking water, sanitation, and hygiene education and promotion

appear in the commitments and investments of both donor governments and international agencies.

The GOP is also a signatory to the Dhaka and Islamabad Declarations on Sanitation. Pakistan's National Sanitation Policy (NSP) envisions the creation of an environment that is free of open defecation, where liquid and solid waste are safely disposed of, and where safe health and hygiene behaviors are practiced in the country (NSP, 2006). NSP objectives include changes in attitudes and behavior regarding the use of sanitation, increased mass awareness of sanitation, and community mobilization.

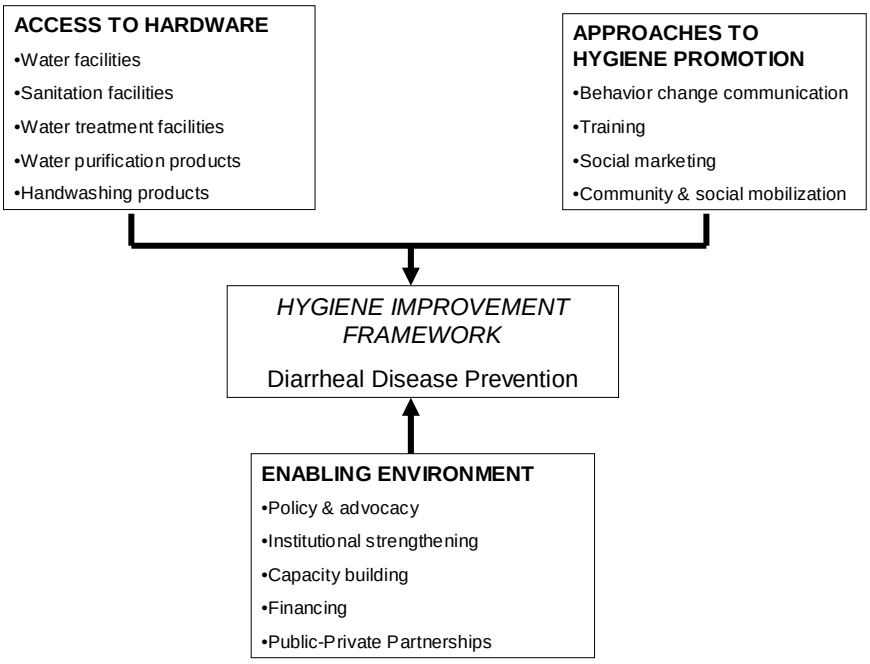
The United States Agency for International Development (USAID) implemented the Pakistan Safe Drinking Water and Hygiene Promotion Project (PSDW-HPP) as part of its goal to improve basic health services for the Pakistani population. The four-year project (2006–2010) was designed to increase the effectiveness and sustainability of the GOP's Clean Drinking Water for All (CDWA) program by conducting complementary hygiene and sanitation promotion programs, community mobilization initiatives, and diverse capacity-building activities. Specifically, the objective of the project was to improve the health of vulnerable populations and to increase the use of proven interventions to prevent waterborne infectious diseases such as diarrhea.

The project was designed to help Pakistan achieve the Millennium Development Goal of a 50-percent reduction in the percentage of its population without access to safe drinking water by 2015. PSDW-HPP's main components included:

- *Hygiene promotion.* Developing, implementing, and evaluating behavior change communications activities to improve safe water management, hand-washing behaviors, and sanitation practices in households. This program included the community hygiene program (including the school program) and the radio program.
- *Capacity building.* Strengthening local capacity to manage and operate CDWA-funded water treatment plants in a sustainable manner.
- *Technical reviews.* Conducting a technical review on water testing methods and household water treatment technologies

Figure 1-1 presents the hygiene improvement framework employed by the project. It details the areas in which intervention is required to ensure reduction and prevention of diarrheal disease and to achieve improvements in hygiene. Under the hygiene promotion component, the project's technical approach targeted mothers and caretakers of children under age five. PSDW-HPP used four key channels (communities, schools, the media, and the private sector) to reach critical audiences on a large scale, delivering behavior change messages and carrying out activities to create sustainable improved hygiene practices.

Figure 1-1: Adapted Hygiene Improvement Framework



PSDW-HPP's geographic scope covered 50 districts and agencies with an estimated total population of 50 million. The project's target areas included four provinces—Sindh, Baluchistan, Punjab, and Khyber Pakhtunkhwa (KPK)—and the Federally Administered Tribal Areas (FATA) and Azad Jammu and Kashmir (AJK), including earthquake-affected areas in KPK, FATA, and AJK.

The initial phase of the project covered 31 districts and agencies with an estimated population of 30 million. These were also the districts where the most intensive hygiene program interventions were implemented in 136 union councils (where filtration plants had been installed). These interventions involved interpersonal communication activities conducted by NGO workers, community theater volunteers, female community volunteers, religious leaders (Maulvis), health units, doctors and other health care providers, and filtration plant operators.

All remaining union councils (UCs) in the project districts and agencies received the project's less-intensive hygiene program and the school hygiene program activities for grade 4 students in government schools (general hygiene program).

The community hygiene program activities – the school activities and both intensive and less intensive hygiene promotion activities were implemented in three phases covering different geographical areas:

- Phase-I (1 Mar 2008 to 14 Mar 2009)
- Phase-II (20 Mar 2008 to 19 Mar 2009)
- Phase-III (01 Jul 2008 to 30 Jun 2009)

Finally, almost all districts in Pakistan—including those that did not receive the community hygiene program—received hygiene promotion and safe drinking water messages through the PSDW-HPP radio campaign. The radio program continued throughout the project implementation period (2006-2010).

This report measures the impact of the community hygiene program, the school program, and the radio campaign in the initial 31 project districts and agencies.

2. Data Collection and Sampling

The baseline survey was conducted in the spring of 2008, and the survey at the end of the project (endline) was conducted in the fall of 2009. Therefore, there was approximately a three month gap between the end of community hygiene activities and the endline survey. The sample for both the baseline and endline surveys comprised mothers or primary caretakers of children between the ages of 0 months and 59 months and that belonged to the lower socio-economic classes—classes C, D, and E as described by ACNielsen and in this paper (see Appendix A for more details). Households were excluded from the survey if they had a fixed water filter or purchased bottled water, as these factors were indicators of a higher socio-economic class. Provinces were treated as a stratum; UCs were the primary sampling units. They were sampled using population proportioned sampling so that no sampling weights were required. Households within the UC were secondary sampling units. The details of the sampling methodology are provided in Appendix B.

In order to collect individual-level panel data, endline survey used the same field methodology as was used in the baseline survey. Given this approach, 51 percent of the households in the baseline sample were captured in the endline survey. A total of 3,934 observations were collected in the baseline and, 3,724 observations in the endline (Table 2-1). Of these, panel information was available for 1,995 households (Table 2-1).

Table 2-1: Distribution of Sample Households by Province and Survey Period

Province Name	Baseline	Endline	Panel
Punjab	1,076	1,078	589
Sindh	1,231	1,232	760
KPK	561	366	191
Baluchistan	504	504	170
AJK	562	544	285
Total	3,934	3,724	1,995

Data sources: 2008 PSDW Baseline Survey, 2009 PSDW Endline Survey

3. Evaluation Design

When estimating the impact of the project on participants, ideally the outcomes (indexed *k* = *1…K)* should be measured after the households participated in the project (y_k^1), and the outcomes, if they had not participated in the project (y_k^0). However, information on the

counterfactual, y_k^0, is never available because either the household participates in a given period, or it does not. Therefore, y_k^0 is often estimated by measuring the outcomes for households that are very similar to the participants but do not participate in the project (non-participants).

The gold standard of measuring project impact involves defining the non-participants using a randomized assignment, and comparing the outcomes across participants and non-participants. These evaluations ensure that that project participation is independent of the outcomes, and the factors influencing them. The average program effect for those subject to random assignment is then estimated as the simple difference in outcomes for those assigned to project and those that are not. However, randomization could not be implemented for this evaluation because the project districts and the project UCs were pre-determined. Once a project district was chosen, the less intensive community hygiene activities were implemented in all the UCs of the project district, and the intensive community hygiene activities were implemented in the UC that had water filtration plants installed by another GOP program. Therefore, there was no flexibility to randomly assign the UCs that would receive the intensive hygiene promotion activities. Furthermore, since the project activities involved community activities so that any household in the UC could participate, it was not possible to randomize participation over the households.

Consequently, the evaluation uses a difference-in-difference (DID) approach to estimate the average program effect on the treated, which requires measuring the key project indicators in the baseline for the control households (y_{1k}^C) and treatment households (y_{1k}^T), and the key project indicators after the project (endline) for the control households (y_{2k}^C) and for the treatment households (y_{2k}^T).[1] The control districts were selected from districts that were similar to the treatment districts as measured by four district-level measures – drinking water access, source of drinking water, respondent's education, and incidence of diarrhea in children under the age of five. Control UCs from these districts were chosen to be in close geographical proximity (although not adjoining UCs to avoid cross over effects of the project), and those that had similar rural/urban status, and similar population density (see Appendix B).

The impact of the project was then evaluated as the DID of key indicators (e.g., percentage of mothers who report washing hands with soap before feeding a child) between the baseline and endline surveys and between the control and treatment households (more on this below). This approach ensures that any systematic differences in the control and treatment households in the baseline are removed when measuring project impact. These systematic differences can be controlled for when DID estimate is derived from parametric estimation by including variables that measure community and household specific characteristics that may be different across control (see equations [4] and [5]).

Thus, the DID estimate for the project impact on outcome *k* is:

$$DID_k = (y_{1k}^T - y_{1k}^C) - (y_{2k}^T - y_{2k}^C) \qquad [1]$$

[1] The subscript 1 denotes baseline and 2 denotes the endline.

The project evaluation was originally intended to measure the impact of the community hygiene program that included the intensive and less intensive community hygiene activities, and the school program. Since the radio programs were being implemented throughout Pakistan, the evaluation was expected to measure only the *incremental impact* of the community hygiene program. Accordingly, as described above, treatment areas were defined as the UCs with the community hygiene program that included the intensive community hygiene activities. The control areas were chosen from non- project districts that did not receive the community hygiene program, other than the nationwide radio programs.

The data from the endline survey suggests that only a very small percentage of households within the treatment area were exposed to the community hygiene program. Given that a very small percentage of the treatment households participated in the community hygiene program, the evaluation plan was modified to measure the outcomes across households that participated in any program –community hygiene program or the radio program -- within the project. That is, the evaluation measured the impact of the treatment on the treated households.

Given the change in the evaluation design, the treatment group was redefined to include those households that participated in any of the PSDW-HPP project components – community hygiene program or the radio program—but did not receive any treatment in the baseline from other programs –the project participants.[2] Since the radio program was nationwide, this implied that some of the project participants were from the original control group (see Table 3-1).

Correspondingly, the control group was redefined to include all households that did not receive any project outputs (including those in the original treatment group who did not participate in any of the projects) –the non-participants. In summary the evaluation measured the impact of the project on households that participated in the program, rather than on households that were *targeted* by PSDW-HPP, some of which may not have participated in the program (average treatment effect.

The project impact was also measured on a subset of the project participants that received the UC-level community hygiene program or school programs to measure the effect of the UC-level activities of PSDW-HPP.

Table 3-1 Evaluation Design: Control and Treatment Households

Treatment and Control Households		
Treatment Households (Participants)	HP[1]/School/Radio Treatment	Households in UCs that participated in *any* program of the project.
	HP/School Treatment	Households in UCs that participated in the community hygiene program
Control Households (Non-Participants)	Households that did not participate in the project.	

Note: [1.] HP – Hygiene Promotion; School – School Program, which together comprise the community hygiene program.

[2] Since the treatment groups were defined based on actual participation, it was not possible to determine if the households participated in USAID's PSDW-HPP program, or some other similar program.

It is important to note that this approach will overstate the overall impact of the project, as the project's impact, is also measured by the reach of the project –the number of households from the target households that participate in the project whether by participating in the community hygiene program activities or by listening to the radio spots. On the other hand, this approach will also understate the results since the non-participants could potentially include households that were geographically or socially close to participants so that the participants could have conveyed the messages to the non-participants leading to some cross contamination.

The average project impact on the participants was measured for the following 18 key project indicators (outcomes):

- Increase in percentage of mothers/caretakers who wash hands with soap after defecation
- Increase in percentage of mothers/caretakers who wash hands with soap after cleaning a child's bottom
- Increase in percentage of mothers/caretakers who wash hands with soap before preparing food
- Increase in percentage of mothers/caretakers who wash hands with soap before eating
- Increase in percentage of mothers/caretakers who wash hands with soap before feeding a child
- Increase in percentage of mothers/caretakers who wash hands with soap at a minimum of two critical times
- Increase in percentage of households with soap located near or inside the kitchen
- Increase in percentage of households with soap located near the toilet
- Increase in percentage of households with soap located at one or more desirable locations (i.e., soap near or inside the kitchen and soap near the toilet)
- Increase in percentage of households covering their drinking water
- Increase in percentage of households storing drinking water in a raised area
- Increase in percentage of percentage of mothers/caretakers that safely take out drinking water
- Increase in percentage of households that treat drinking water using project-promoted methods
- Decrease in percentage of percentage of mothers/caretakers with the belief that clear, odorless water is safe to drink
- Increase in percentage of percentage of mothers/caretakers that air dry their hands.

3.1.1 Individual Panel Data Difference-in-Difference Estimator

DID estimation using individual-level panel data is a powerful tool for estimating policy effects, where project impact for household, *i,* is measured as the change in outcome *k* in the baseline (y_{i1k}), and outcome *k* in the endline period (y_{i2k}), where 1 denotes baseline and 2 denotes the

endline, and is equal to $\Delta y_{ik} = y_{i1k} - y_{i2k}$.[3] This is the dependent variable in the estimation equation and is defined as a dummy equal to 1 if the household exhibited a positive change in the outcome over the endline and baseline and zero otherwise.[4] The project impact for outcome *k* is then estimated using a difference-in-difference estimation given as:

$$\Delta y_{ik} = (y_{i2k} - y_{i1k}) = \eta_k + \mu_k (dT_{i2} - dT_{i1}) + \delta_k (dTV_{k2} - dTV_{k1}) + (u_{i2k} - u_{i1k}), \qquad [2]$$

i = 1…N for all *k* = 1….18 outcomes

where N is the number of households, *k* are the 18 different outcomes mentioned in the section above. The variable dTV_{it} is a dummy variable equal to 1 if household viewed a TV spot as part of another program. The dummy variable dT_{it} is 1 if the household participated in the project in the period *t* =*1,* 2 (baseline and endline) and zero otherwise, which captures the difference in outcome *k* between the participants and non-participants prior to the project. In this case, since all households participate in the project only in the endline, dT_{i1}=0, so that the term is simply dT_{i2}, which is a dummy equal to 1 if the household participated in the project in the endline period, and u_{itk} are the idiosyncratic errors. In a linear regression, OLS estimator, $\hat{\mu}_k = \Delta\bar{y}_k^T - \Delta\bar{y}_k^C$ is the DID estimate where the difference is taken across time periods for the same households. However, since the outcome variable is binary, equation [2] is estimated using a non-linear probit estimation and the DID estimate is estimated as[5]:

$$DID_k^{Panel.Model\ 1} = \Phi(\hat{\eta}_k + \hat{\mu}_k dT_2 + \hat{\beta}_k \overline{dTV}_k) - \Phi(\hat{\eta}_k + \hat{\beta}_k \overline{dTV}_k) \qquad [3]$$

where Φ is the cumulative normal density function, and the bar denotes variable means.

In another variation of equation [2] (Model 2), *dT* is a vector of all specific program treatments (dT = {dHP, dRadio}) so that the impact of the community hygiene program and radio program can be measured separately (see Table 3-1). Therefore, the DID panel estimate for model 2 is:

$$\begin{aligned} DID_k^{Panel\ Model\ 2} = {} & \Phi(\hat{\eta}_k + \hat{\mu}_{1k} dHP_2 + \hat{\mu}_{2k} \overline{dRadio}_2 + \hat{\beta}_k \overline{dTV}_k) \\ & - \Phi(\hat{\eta}_k + \hat{\mu}_{2k} \overline{dRadio}_2 + \hat{\beta}_k \overline{dTV}_k) \end{aligned} \qquad [4]$$

As described in Table 2-1, panel information was available for only 51 percent of the households that were interviewed in the baseline. Even though the panel DID estimate is based on a smaller number of observations, it is useful to look at the impact on the cohort of individuals for whom

[3] In the panel data, the outcome for the same household is compared over time; therefore there is no need to take a difference across the control and treatment.

[4] Households that either had a desirable outcome in the endline and the baseline, or had a positive outcome in the baseline but not in the endline, were assigned a dummy equal to zero. Only households that did not have desirable outcome in the baseline, and had a positive outcome in the endline were assigned a dummy equal to one.

[5] Instead of estimating the DID estimate as the average marginal effect at the mean of other explanatory variables (dTV, dRadio), we could estimate it as the average of the estimated predicted probabilities at each observation. However, the difference in the estimates using this procedure was negligible.

data was available from the baseline and endline. Particularly, because this analysis controls for all household and individual specific differences, and provides more robust estimates.

The concern with the panel data analysis is two-fold. First, the inference can be different insofar as the sample of households is smaller than that included in the pooled analysis. Second, the concern can be that the if attrition of households (loss of baseline households in the endline) is systematic so that only certain types of households are dropped in the endline, or move away from the area, then our results can be biased.

However, at the outset, it was not expected that a certain type of households are less or more likely to be not available for a response in the endline, or have moved out of the area. The only remote possibility could have been that the households that responded in the baseline and did not improve their hand washing and sanitation practices may have preferred to not respond in the endline rather than admit to not changing their behavior. However, a large part of the reason information on the baseline household could not be collected was either because the household had moved away, or because the endline survey relied on a less than perfect approach to capture the baseline households.[6] This means that there is no reason to expect that the attrition rate is not random.

3.1.2 Difference-in-Difference Pooled Data

In addition to the panel data analysis, the project impact was estimated by combining all the households in a pooled sample, disregarding the panel information. This pooled data included information on additional 1,724 households for whom only endline data were available. Because treatment groups were defined based on the endline information on whether the household participated in the project, households for whom there was no endline data could not be included in the pooled analysis.

The model for estimating DID estimate for each outcome *k*, is:

$$y_{ik} = \alpha_k + \boldsymbol{\beta}_k \mathbf{X}_k + \delta_k dT_{ik} + \gamma_k d2_{ik} + \mu_k dT_{ik}.d2_{ik} + u_{ik}, \qquad [4]$$

i = 1….N for all *i* = 1,…18

where N is the number of households, y_{ik} is the dummy variable for outcome *k*, and is equal to one if the outcome is desirable, and zero otherwise, X_k is a matrix of covariates, which would control for compositional changes across the participants and non-participants, and where relevant over the change over time. These covariates include mother's years of education, which is expected to have an impact on the outcomes, the size of the household, number of children, a composite measure of household assets, and viewing of TV spots on hand washing that were delivered by other GOP projects (see Table 3-2).[7] In addition to these variables, a dummy for the province, and the SEC category of the households was also included.

[6] The endline survey was conducted in the baseline areas, and the field strategy mimicked that in the endline where the interviewers started from the same starting point, and used the same strategy to interview the households (see Appendix A for details.)

[7] Using the information collected in the endline as dummy variables that were equal to one if the household owned assets such as livestock, consumer durables (TV, fridge etc.), vehicles, an asset index was created using principal component and factor analysis (Donnel et al, 2007). Ideally, it is preferable to

The time dummy $d2_{ik}$, is equal to 1 if the time period is endline, and zero otherwise. For example, suppose the first outcome is whether the household cover their water containers with a hardcover. The dummy variable y_{i1}, will be 1 if the household covers their container with a hardcover, and zero otherwise. Therefore, the coefficient on this dummy, γ_k, captures the aggregate factors that would cause changes in outcome y_{ik} even in the absence of the project. The coefficient μ_k is the difference-in-difference estimate.[8] As before, in the case of binary outcomes, a probit specification is used to estimate the model. The coefficient δ_k is the DID pooled estimate for outcome *k* and is equal to:

$$DID^{Pool} = \hat{\delta}_k = [\Phi(\hat{\alpha}_k + \hat{\boldsymbol{\beta}}_k \overline{\mathbf{X}}_k + \hat{\delta}_k \overline{dT} + \hat{\gamma}_k \overline{d2} + \hat{\mu}_k) - \Phi(\hat{\alpha}_k + \hat{\boldsymbol{\beta}}_k \overline{\mathbf{X}}_k + \hat{\delta}_k \overline{dT} + \hat{\gamma}_k \overline{d2})] \quad [5]$$

where Φ is the predicted probability from probit regression. The significance levels are based on linear regression of the same model.

Table 3-2 Summary Statistics of Key Household Characteristics

Variable	Baseline				Endline			
	Control		Treatment		Control		Treatment	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Number of Children	4.0	2.2	4.1	2.5	3.7	1.9	3.9	2.0
Mother's Education (years)	0.9	2.8	4.7	5.1	3.2	4.4	4.5	5.0
Dummy =1 if Household has Flush Toilet	0.71	0.46	0.64	0.48	0.79	0.40	0.86	0.35
Asset Index[1]	-1.6E-03	0.5	0.1	0.9	-4.0E-02	1.0	0.2	0.9
Household Size	7.6	3.1	7.6	3.5	7.9	3.6	8.1	3.9
Viewed TV Spots on Hygiene	0.3	0.5	0.8	0.6	0.5	0.6	0.7	0.5

Data sources: 2008 PSDW Baseline Survey, 2009 PSDW Endline Survey

Notes: [1.] Using the information collected in the endline on the ownership of assets such as livestock, consumer durables (TV, fridge etc.), vehicles, an asset index was created using principal component and factor analysis (Donnel et al, 2007).

use information on baseline, but these data were collected only in the endline. In this case, using endline information on household assets is not a concern for two reason. First, endline information was collected within three months of the baseline, which is a short period of time for household asset holdings to change. Second, the project was not expected to impact the household asset situation, so that there is no concern about endogeneity.

[8] Ai and Norton (2006) note that the interaction term in non-linear regressions such as the probit may not be a reliable estimator of the true interaction effect. However, Puhani (2008) notes, the coefficient of the interaction effect is the true treatment effect (difference-in-difference estimator).

4. Results

Section 4.1 presents the basic socio-economic characteristics of the sample, section 4.2 presents the results of the evaluation that include DID estimates for Model 1 and Model 2 using the panel and pooled data.

4.1 Basic Characteristics

All the results are presented for the pooled data, by project participants (the control group) and non-participants (the treatment group) as defined above. The data in all tables are from the baseline and endline household survey conducted by ACNielsen.

Based on GOP criteria, urban and rural households were classified into socio-economic classes (SECs). Urban households were classified into five SECs based on the education and occupation of the chief wage earner, while the rural households were classified based on the structure of the house and education level of the head of the household. Appendix A contains a more detailed description of the classifications.[9]

For the rural households, the classification was based on four types of houses—*pukka* upper, *pukka* lower, semi-*pukka,* and *kuchha*[10]—and six levels of education, ranging from illiterate to post-graduate.[11] Details on the types of houses and the level of education are provided in more detail in Appendix A. Rural households were classified into five SEC categories from the highest (A) to the lowest (E), based on the following five criteria:

- The education level of the head of household is at least intermediate, and the structure of the house is either *pukka* lower or *pukka* upper.
- The education level of the head of household is up to matriculation level (10 years of education), and the structure of the house is any of the four types.
- The education level of the head of household is less than matriculation level, and the structure of the house is any of the four types. This is the middle class of rural Pakistan.
- The head of household is illiterate (which is very common). The structure of the house is either semi-*pukka* or *pukka* lower.
- The head of household has no formal education. The structure of the house is *kuchha.*

Urban households were divided into five SECs based on the education of the head of household and the occupation of the chief earner. The households were classified based on 11 occupational categories and, as with the rural areas, education categories (see Appendix A for a detailed description of the occupational categories). Urban households were classified into five SEC categories from the highest (A) to the lowest (E), based on the following criteria:

[9] The program was intended to focus on the SEC categories C, D and E. However, based on the response by the households, some of the households appear to be in SEC categories A, B, and C also. It is possible that some of this may be because their status changed over the baseline and end-line analysis.

[10] Kuccha means houses that are not very permanent and are made of mud or similar materials. Pukka means permanent houses made of concrete, bricks, or other durable materials.

[11] Illiterate, up to primary (less than 5 years), matriculation (6-9 years of school), intermediate (10 years), graduate (12 years), and post-graduate.

- Well-educated, self-employed or employed professionals, senior-level executives or officers in public or private limited organizations, well-educated small- to large-scale businesspeople and supervisors
- Relatively less well-educated lower or mid-level executives and officers, well-educated small businesspeople and supervisors
- Predominantly small retailers and businesspeople, supervisors, and lower-level executives who have 5-10 years of schooling
- Relatively well-educated skilled workers, less well-educated small retailers, and non-executive staff members
- Skilled or unskilled workers, petty traders, and non-executive staff members with no more than 10 years of schooling

Table 4-1, below shows the breakdown of household SECs by province in the pooled data. The majority of the households in each province were classified as belonging to class E. Baluchistan had the highest proportion of households in class E, while AJK had the lowest. Correspondingly, Baluchistan had the lowest percentage of households in class C, while AJK had the highest. SEC distributions in Punjab and KPK were very close (within 3.5 percent at the most), while Sindh had a greater-than-average proportion of class E households.

Table 4-1 Distribution of Households by Province and Socio-Economic Class

Province	Baseline Socio-Economic Class						Endline Socio-Economic Class					
	A	B	C	D	E	Total	A	B	C	D	E	Total
Punjab		1%	5%	33%	61%	949	1%	3%	10%	26%	60%	935
Sindh				29%	71%	1121		3%	5%	27%	64%	1121
KPK			2%	37%	62%	511	4%	6%	5%	26%	59%	312
Baluchistan				22%	78%	479	1%	1%	1%	20%	76%	471
AJK	3%	8%	15%	22%	52%	432	4%	13%	19%	19%	45%	467
Total	0%	1%	4%	29%	66%	3492	0%	5%	8%	26%	65%	3306

Data sources: 2008 PSDW Baseline Survey, 2009 PSDW Endline Survey

As illustrated in Table 4-2, approximately 54 percent of the participants were in class E (the lowest SEC), while 63 percent of the sample in the control group is in class E. Among the provinces, Baluchistan had the largest percentage of non-participants in the lowest SEC, with as many as 76 percent of the non-participants in this class (73 percent for participants).

Table 4-2 Distribution of Households by Province, Socio-Economic Class

Province Name	Non-Participants						Participants					
	Socio-Economic Class						Socio-Economic Class					
	A	B	C	D	E	Total	A	B	C	D	E	Total
Punjab	1%	3%	10%	25%	61%	826	2%	2%	12%	30%	54%	109
Sindh		3%	5%	26%	66%	928		5%	5%	35%	55%	193
KPK	4%	5%	4%	26%	61%	250	5%	10%	10%	24%	52%	62
Baluchistan	1%	1%	2%	20%	76%	422		2%		24%	73%	49
AJK	4%	12%	19%	19%	46%	364	5%	14%	20%	18%	43%	103
Total	1%	4%	8%	24%	63%	2790	2%	6%	10%	28%	54%	516

Data sources: 2009 PSDW Endline Survey

4.2 Drinking Water Access and Treatment

The baseline data indicate that among the participants, 66 percent of the households had their main source of drinking water inside the house, while 76 percent of non-participants had their main source of drinking water inside the house. The remaining households had their main source of drinking water outside the house (Table 4-3). In the endline survey, among the non-participating households, the proportion of households that got their drinking water inside the home remained the same (76 percent), and among the participants, there was a slight increase in inside sources of water to 76 percent.

Table 4-3: Distribution of Households by Main Source of Drinking Water

	Baseline				Endline			
	Non-Participants[12]		Participants		Non-Participants[13]		Participants	
Province Name	Inside	Total	Inside	Total	Inside	Total	Inside	Total
Punjab	84%	1,020	88%	49	83%	1,020	80%	132
Sindh	85%	1,184	63%	32	85%	1,184	88%	221
KPK	77%	507	54%	37	88%	507	90%	71
Baluchistan	61%	472	84%	31	60%	472	73%	52
AJK	48%	481	46%	54	51%	481	43%	120
Total	76%	3,664	66%	203	76%	3,664	76%	596

Data sources: 2008 PSDW Baseline Survey, 2009 PSDW Endline Survey

The World Health Organization defines *improved* drinking water sources as tap water inside the house (household connections) and outside the house (public standpipes), protected dug well, boreholes (includes tube wells), protected springs, and rainwater collection (WHO 2007). All other sources of water such as unprotected well, unprotected springs are considered unsafe. It is important to note that these sources only reflect improved water sources and do not guarantee safe water (WHO 2007). Table 4-4 presents the main sources of improved drinking water inside the house for households that reported that their main source of drinking water was inside the house.

The results suggest that a large majority of the households had access to improved sources of water inside the house, and there was no real increase in the percentage of households with access to improved sources from the baseline to endline. Among the participants, the percentage of households that reported having access to drinking water reduced from 95 percent to 99 percent over the analysis period. And this percentage remained the same for the non-participants (99 percent). This suggests that if the household had access to drinking water inside the house, it was largely from an improved source.

[12] When "HP" or "TV/Radio" isn't specified, the control/treatment variables refer to the total treatment variable which includes both the HP and TV/Radio groups (i.e., if a household either participated in the hygiene program *or* was exposed to a TV/Radio spot, it is counted as "treatment").

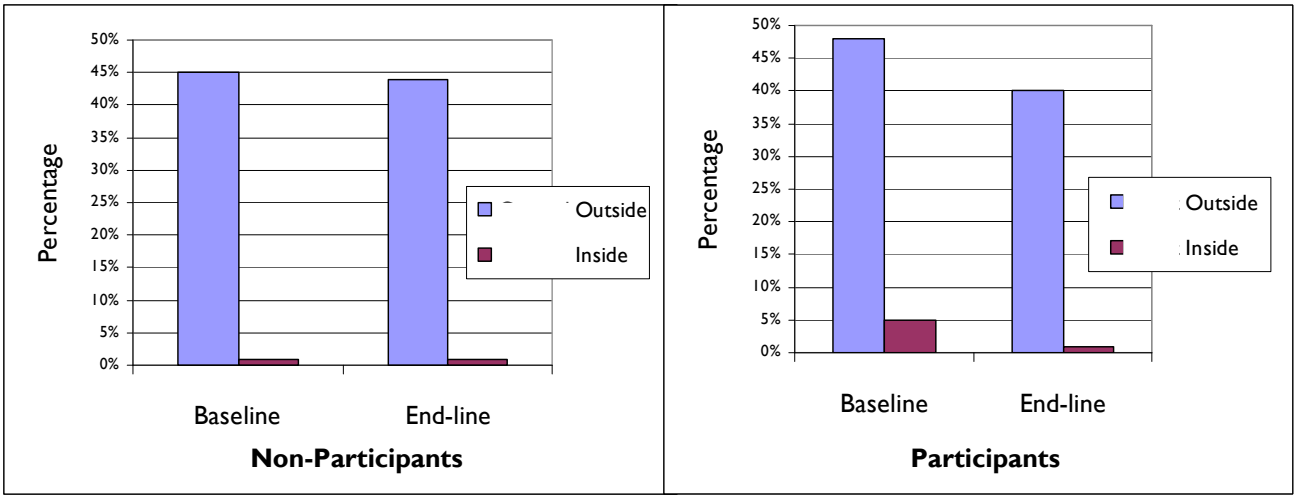
[13] When "HP" or "TV/Radio" isn't specified, the control/treatment variables refer to the total treatment variable which includes both the HP and TV/Radio groups (i.e., if a household either participated in the hygiene program *or* was exposed to a TV/Radio spot, it is counted as "treatment").

Table 4-4: Distribution of Households with Improved Main Drinking Water Source Inside the House

	Province Name	Non-Participants				Participants			
		Unsafe Outside Source		Unsafe Inside Source		Unsafe Outside Source		Unsafe Inside Source	
		%	Total	%	Total	%	Total	%	Total
Baseline	Punjab	40%	164	0%	1,020	33%	6	0%	49
	Sindh	37%	176	0%	1,184	75%	12	0%	32
	KPK	35%	116	2%	507	29%	17	10%	37
	Baluchistan	48%	184	6%	472	40%	5	15%	31
	AJK	57%	252	3%	481	52%	29	4%	54
	Total	45%	892	1%	3,664	48%	69	5%	203
End-line	Punjab	29%	162	1%	1,020	23%	26	0%	132
	Sindh	41%	152	0%	1,184	31%	26	0%	221
	KPK	28%	36	5%	507	43%	7	0%	71
	Baluchistan	60%	182	0%	472	64%	14	0%	52
	AJK	47%	206	2%	481	45%	69	6%	120
	Total	44%	738	1%	3,664	40%	142	1%	596

Note: The rows may not add to total because of rounding.
Data sources: 2008 PSDW Baseline Survey, 2009 PSDW End-line Survey

Figure 4-1: Distribution of Households by Unsafe Outside and Inside Drinking Water Source



. Among the nearly quarter of participating households that got their drinking water from an outside source, the percentage of unsafe outside sources reduced from 48 percent in the baseline to 40 percent in the endline. This is compared to only a one percent reduction among the non-participants, from 45 percent to 44 percent.

An important factor that affects hygiene and sanitation is the access to toilets. There was a dramatic reduction among households without access to any toilets among the participants, from 28 percent in the baseline to 8 percent in the endline. There was a smaller reduction among the non-participants, from 19 to 12 percent.

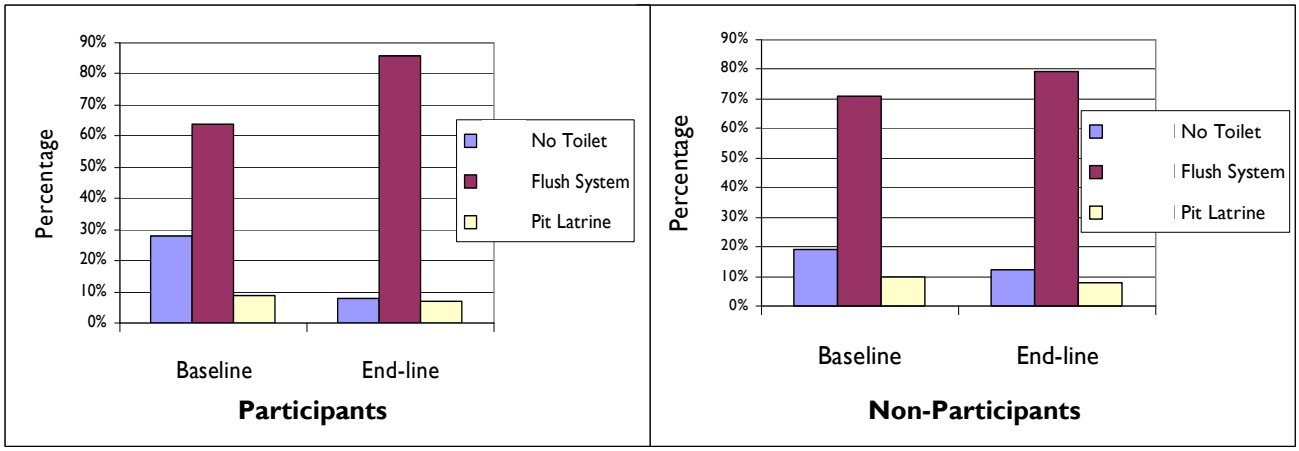
Among the participating households with access to a flush system increased from 64 percent in the baseline to 86 percent in the endline. This is compared to 71 percent and 79 percent among the non-participants (see Table 4-5).

Table 4-5: Distribution of Households with Access to Toilet

	Province Name	Non-Participants				Participants			
		No Toilet	Flush System	Pit Latrine	Total	No Toilet	Flush System	Pit Latrine	Total
Baseline	Punjab	6%	89%	5%	1,020	0%	98%	2%	49
	Sindh	31%	60%	8%	1,184	47%	34%	19%	32
	KPK	13%	84%	2%	507	35%	62%	3%	37
	Baluchistan	22%	39%	38%	472	32%	42%	26%	31
	AJK	19%	77%	4%	481	33%	63%	4%	54
	Total	19%	71%	10%	3,664	28%	64%	9%	203
Endline	Punjab	4%	91%	5%	946	7%	90%	3%	49
	Sindh	9%	83%	8%	1011	5%	92%	4%	32
	KPK	3%	88%	9%	295	4%	94%	1%	37
	Baluchistan	39%	38%	23%	452	15%	48%	37%	31
	AJK	15%	82%	3%	424	13%	80%	8%	54
	Total	12%	79%	8%	3128	8%	86%	7%	203

Data sources: 2008 PSDW Baseline Survey, 2009 PSDW Endline Survey

Figure 4-2 Distribution of Households by Access to Toilet



4.3 PSDW-HPP Project Impact

The USAID PSDW-HPP project intended to affect several aspects of sanitation and hygiene behavior that included three broad areas:

1) Drinking water storage
2) Hygiene practices
3) Knowledge and treatment of water

Accordingly, the impact evaluation assessed the extent to which the project changed the indicators that measured these outcomes. The results of the evaluation are presented for both the DID estimate using the pooled data, and the DID estimate where the analysis was limited to individual panel data.

The project impacts on the three major components of the project are discussed in the next few sections.

4.3.1 Impact on Drinking Water Storage

One of the areas in which the project intended to make a change is how households store their drinking water. Therefore, this aspect of the project was relevant to the households that stored water in containers. The data suggests that approximately 69 percent of the non-participants stored water in containers in the baseline, and this percentage increased to 81 percent in the endline (Table 4-6). Among the participants this percentage increased from 69 percent in the baseline to 83 percent in the endline.

Table 4-6: Distribution of Households by Drinking Water Storage

	Province Name	Non-Participants				Participants			
		Containers	Roof	No Storage	Total	Containers	Roof	No Storage	Total
Baseline	Punjab	55%	18%	27%	1,020	73%	12%	14%	49
	Sindh	87%	4%	8%	1,184	63%	16%	22%	32
	KPK	76%	14%	10%	507	84%	11%	5%	37
	Baluchistan	27%	52%	21%	472	19%	52%	29%	31
	AJK	85%	10%	5%	481	87%	9%	4%	54
	Total	69%	16%	15%	3,664	69%	18%	13%	203
Endline	Punjab	75%	10%	15%	942	75%	10%	15%	129
	Sindh	90%	6%	4%	1011	90%	7%	4%	221
	KPK	74%	15%	12%	295	76%	14%	10%	71
	Baluchistan	70%	25%	5%	442	65%	30%	4%	46
	AJK	89%	8%	3%	424	89%	6%	5%	120
	Total	81%	11%	8%	3114	83%	10%	7%	587

Data sources: 2008 PSDW Baseline Survey, 2009 PSDW Endline Survey

The project's community hygiene program, school and radio program promoted the following behaviors to households that store drinking water in containers:

- Store the drinking water by covering the containers with a hard cover
- Storing the containers in a raised area
- Take out water safely from the containers either using a long handled scoop or using containers that have a tap

Among the households that store their drinking water in containers, the project's objective was to increase the percentage of households that store their drinking water safely. Table 4-7 presents the impact evaluation results for the water storage indicators using pooled data. The table presents the percentage of households by drinking water storage behaviors in the baseline, after the program for both program participants and non-participants. On average, in the baseline, a larger percentage of households covered their drinking water containers with a hardcover if they participated in the project. After the project, in the endline, *both* participants and non-participants reported a large percentage increase in this indicator. Since non-participants also experienced an increase in the indicators, the changes can't be attributed exclusively to the project interventions.

Among the households that participated in the project, the percentage of households using a hard cover on their drinking water containers increased from 59 percent to 75 percent compared with an increase of 56 percent to 72 percent for households in the control area (Table 4-7). The last column presents the DID pooled estimate, which is positive but insignificant for all storage behaviors, implying that the improved storage behavior cannot be attributed to the PSDW-HPP project.[14]

[14] The DID estimate reported in the table is not equal to the simple difference in percentage across participants and non-participants before and after the program, and is estimated using equation [5], which controls for community and household characteristics that might impact the outcome.

Table 4-7: Impact of USAID-PSDW-HPP on Storage of Water Containers

PSDW-HPP Project Impact	Baseline		Endline		DID Pooled Estimate[1]
Storage Indicators	Non-Participants	Participants	Non-Participants	Participants	
Covered with Hard Cover	56%	59%	72%	75%	4% 0.45)
Stored in a Raised Area	82%	81%	89%	91%	1% (0.03)
Safe Method Take Out Water	55%	63%	51%	50%	-3% (-0.48)

* Significant at the 10% level
** Significant at the 5% level

Table 4-8 presents the results of the pooled and panel data estimation side by side. Panel data estimate also includes the impact of the project on the households that participated in the community hygiene program (separate from households that only participated in the radio program). It is possible that households that participated in the community hygiene program have better outcomes because these programs reinforced radio messages, and were conducted in closer contact with the community. However, the DID panel data estimates do not suggest additional positive impact among the households that participated in the community hygiene program.. [15]

Table 4-8: Impact of USAID-PSDW-HPP on Storage of Water Containers

PSDW-HPP Project Impact		Covered with Hard Cover	Stored in a Raised Area	Safe Method to Take Out Water
DID, Pooled Data				
Model 1	HP/School/Radio	0.04	0.01	-0.03
		(0.45)	(0. 03)	(-0.48)
DID, Panel Data				
Model 1	HP/School/Radio	0.01	-0.01	-0.02
		(0.21)	(-0.34)	(-0.73)
Model 2	HP/School	-0.03	0.01	-0.01
		(-0.57)	(0.30)	(-0.34)
	Radio Only	0.07	-0.04	-0.03
		(1.13)	(-1.04)	(-0.82)

* Significant at the 10% level
** Significant at the 5% level

4.3.2 Impact on Hygiene Practices

Another important objective of the PSDW-HPP project was to create positive behavior change in hand washing practices of mothers/caretakers of children at five critical times:

1) After cleaning a child's bottom,

2) After defecating

[15] The possibility that the program could have different impact across the provinces was also considered. However, the data was not adequate to estimate province-specific impact reliably.

3) Before preparing food

4) Before feeding a child

5) Before eating

In the baseline, 68 percent of the mothers/caretakers that participated in the project reported washing hands with soap after defecating compared with 92 percent in the endline. This compares to a change of 74 percent to 91 percent among the non-participants. The DID estimate using pooled data suggest that households that participated in any of the project activities had a positive but a marginally significant change in the practice of washing hands after defecating (Table 4-9).

While there were dramatic changes in the other promoted hygiene behaviors also between baseline and endline, these were observed by both participants and non-participants. Although, the DID estimates a positive but statistically insignificant impact of the project on these behaviors. Therefore, the improvement in hygiene behaviors cannot be attributed to the PSDW-HPP project. This could be because the radio programs were nationwide and even though households reported not listening to the radio program, they may have received similar messages from their neighbors and community members that listened to those messages. There is no way to isolate the spillover effects but it is a potential explanation why there was improvement in hand washing behavior of both project participants and non-participants so that the increase cannot be attributed to the PSDW-HPP project.

Table 4-9: Impact of USAID-PSDW-HPP on Hand Washing Practices

PSDW-HPP Project Impact	Baseline		Endline		DID Pooled Estimate
Handwashing Indicators	Non-Participants	Participants	Non-Participants	Participants	
Hand Washing after Cleaning Child's Bottom	73%	64%	81%	77%	2% (0.80)
Hand Washing after Defecating	74%	68%	91%	92%	3% (1.42)
Hand Washing before Preparing Food	62%	62%	72%	72%	-2 (-0.32)
Hand Washing before Feeding Child	54%	53%	74%	77%	3% (0.50)
Hand Washing before Eating	46%	44%	74%	75%	3% (0.58)

** Significant at the 5% level

Table 4-10 presents the DID estimate for panel data, and also the impact of participation in the community hygiene program. The panel data analysis finds a negative impact of the project on hand washing behavior. In fact the impact is negative and significant on hand washing after defecating.

Self reporting of hand washing behavior is known to be unreliable because of the proclivity of individuals to over report positive behavior. Thus, it is possible that self-reported nature of the data is leading to some measurement error. The results on the next set of outcomes are based on direct observations: whether households keep soap in the appropriate location for hand washing.

Table 4-10: Impact of USAID-PSDW-HPP on Critical Hand Washing Behavior

PSDW-HPP Project Impact		Washing after Cleaning Child's Bottom	Washing after Defecating	Washing before Preparing Food	Washing before Feeding Child	Washing before Eating
DID, Pooled Data						
Model 1	HP/School/Radio	0.02	0.03	-0.02	0.03	0.03
		(0.80)	(1.42)	(-0.32)	(0.50)	(0.58)
DID, Panel Data						
Model 1	HP/School/Radio	-0.06**	-0.05*	0.04	-0.04	-0.04
		(-2.11)	(-1.74)	(1.44)	(-1.22)	(-1.20)
Model 2	HP/School	-0.07**	-0.05	0.04	-0.03	-0.07
		(-2.09)	(-1.46)	(1.21)	(-0.83)	(-1.64)
	Radio Only	-0.04	-0.05	0.04	-0.05	0.01
		(-0.80)	(-1.03)	(0.86)	(-1.01)	(0.21)

* Significant at the 10% level
** Significant at the 5% level

As part of the survey, the interviewers observed where the soap was kept in the household and if there was soap in desirable locations (i.e., soap near or inside the kitchen and soap near or inside the toilet). Since this information was collected by direct observation, it is expected to have minimal measurement error.[16] The results presented in Table 4-11 indicate that participation in any of the project activities had a positive and significant impact on the practice of keeping soap near the toilet (9 percent increase) and the practice of keeping soap near the kitchen (5 percent increase). As expected, the PSDW-HPP project also had a positive and significant impact on the practice of keeping a soap in at least the kitchen or the toilet (12 percent increase).

According to the panel data DID estimate, the impact of the PSDW-HPP project was positive and significant on the practice of location soap near the kitchen (7 percent increase) and a marginally significant impact on the practice of locating soap near the toilet. The panel data DID estimate suggests that participation in the community hygiene program had a positive and significant impact on almost all hygiene practices: the practice of locating a soap near the toilet (6 percent increase), the practice of locating soap near the kitchen (8 percent increase), and the practice of locating soap in at least one location (7 percent increase).

[16] A possible source of measurement error in this variable could be the variation among interviewers on what the definition of "near" is. It is also possible that over the baseline and endline there is a difference in the definition where the training modules are different. There is some concern about this variable that the definition of "near" in the endline was less stringent, which could overestimate program impact.

The evaluation also estimated the impact of the project on the percentage of mothers or caretakers who reported washing hands with soap at least two out of the five critical times (Table 4-11). The DID pooled data results find only a marginally significant impact of the project. Consistent with the results above, the DID estimate using panel data suggests that the project did not have a positive impact on hand washing at two or more critical times. Surprisingly, project participation appears to have a negative impact on this indicator.

Table 4-11: Impact of USAID-PSDW-HPP on Key Hygiene Practices

PSDW-HPP Project Impact	Baseline		Endline		DID Pooled Estimate
Hygiene Indicators	Non-Participants	Participants	Non-Participants	Participants	
Hand Washing at least of Two Critical Times	73%	64%	81%	77%	2%
					(1.28)
Soap near Kitchen	5%	4%	10%	13%	5%*
					(1.66)
Soap near Toilet	49%	42%	61%	64%	9%*
					(1.81)
Soap in one or more locations	52%	44%	65%	70%	12%**
					(2.36)
Air Drying of Hands	11%	11%	10%	15%	4%
					(1.20)

* Significant at the 10% level
** Significant at the 5% level

Even after proper washing of hands with soap, there can be a chance of contamination if hands are not dried in a safe manner (either using a clean towel or air drying). Although using a clean towel would normally be considered safe, the definition of "clean" can be confounded by a cloth that looks clean but is not actually clean or by a different understanding of what clean actually is. The DID estimate using pooled data suggests that participation in any of the PSDW-HPP programs led to an increase in the percentage of mothers and caretakers that reported air drying their hands (3 percent increase), although the impact was shy of being significant at the 10 percent level. The DID estimate using panel data suggests a higher magnitude of impact (9 percent increase), and also finds that participation in the community hygiene program had a positive and significant impact on the practice of air drying hands (Table 4-12).

Table 4-12: Impact of USAID-PSDW-HPP on Critical Hand Washing Behavior

PSDW-HPP Project Impact		Hand washing at least Two Critical Times	Soap Near Kitchen	Soap Near Toilet	Soap in more than one Locations	Air Drying of Hand
DID, Pooled Data						
Model 1	HP/School/Radio	0.02	0.05*	0.9*	0.12**	0.04
		(1.28)	(1.69)	(1.81)	(2.36)	(1.20)
DID, Panel Data						
Model 1	HP/School/Radio	-0.06**	0.07**	0.04	0.05*	0.09**
		(-1.98)	(3.48)	(1.34)	(1.70)	(4.56)
Model 2	HP/School	-0.07**	0.08**	0.06*	0.07**	0.07**
		(-2.07)	(3.30)	(1.66)	(2.01)	(3.04)
	Radio Only	-0.03	0.05	0.001	0.01	0.13**
		(-0.59	(1.55)	(0.01)	(0.14)	(3.84)

* Significant at the 10% level
** Significant at the 5% level

4.4 Knowledge of and Attitudes toward Safe Drinking Water

An important component of the USAID PSDW-HPP project was to increase the knowledge and attitude toward safe drinking water, and the proper treatment of drinking water. The results suggest that across the baseline and endline, the percentage of households that reported positive behavior change with respect to these indicators increased, but that increase was observed among the participants and non-participants suggesting, similar to other indicators, that the increase was not attributable to the project. Consequently, the DID pooled data estimates suggest that the project did not have a statistically significant impact on knowledge of safe drinking water, perception that clear water is safe to drink, or increase the percentage of mothers/caretakers that understood the correct reasons for diarrhea. (Table 4-13)

Table 4-13: Impact of USAID-PSDW-HPP on Treatment of Water and Knowledge

PSDW-HPP Project Impact	Baseline		Endline		DID Pooled Estimate
Storage Indicators	Non-Participants	Participants	Non-Participants	Participants	
Correct Treatment of Drinking Water	47%	44%	57%	58%	1%
					(0.03)
Belief that Clear Water is Safe to Drink	70%	69%	86%	83%	-5%
					(-1.05)
Knowledge about the correct reasons for Diarrhea	54%	57%	65%	70%	1%
					(0.21)

Interestingly, the DID panel data estimate, finds that the impact of the PSDW-HPP project was significant on the percentage of households that reported treating their water correctly. The results indicate that the project led to a 4 percent increase in the percentage of households that reported treating their water correctly (Table 4-14).

Table 4-14: Impact of USAID-PSDW-HPP on Treatment of Water and Knowledge

PSDW-HPP Project Impact		Treatment of Water	Perception that Clear Water Safe to Drink	Correct Knowledge about Reasons for Diarrhea
DID, Pooled Data				
Model 1	HP/School/Radio	0.01	-0.05	0.01
		(0.03)	(-1.05)	(0.21)
DID, Panel Data				
Model 1	HP/School/Radio	0.04**	0.03	-0.04
		(2.80)	(1.04)	(-1.41)
Model 2	HP/School	0.04**	0.03	-0.05
		(2.18)	(0.95)	(-1.28)
	Radio Only	0.05**	0.03	-0.04
		(1.97)	(0.50)	(-0.70)

** Significant at the 5% level

5. Summary Discussion

The final impact evaluation of the USAID PSDW-HPP used a DID approach on both pooled and panel data to estimate program impact. Both the DID pooled and DID panel data estimates find that the improvement in the hand washing behavior at critical times cannot be attributed to the PSDW-HPP project. Self reporting of hand washing behavior is known to be unreliable because of the proclivity of individuals to over report positive behavior. Thus, it is possible that self-reported nature of the data is leading to some measurement error.

The results on outcomes that are based on direct observations were more promising: whether households keep soap in the appropriate location for hand washing. The evaluation finds that the project had a positive and significant impact on the practice of locating soap in at least one location, either near a kitchen or a toilet. For hand washing after defecating the evidence was conflicting across the DID pooled data estimate (positive and marginally significant) and panel data estimates (negative impact).

The panel data estimate finds that participation in the project led to a 4 percent increase in the percentage of households that treated their drinking water correctly, which was statistically significant. The panel data estimate also finds a positive impact of the project on the practice of air drying hands after washing.

In general, the magnitude of impact varied across households that participated in any program component and households that participated in the community hygiene program. Except in the case of air drying of hands, if the impact of the project on an indicator was positive and statistically significant, the magnitude of the impact was greater among the households that participated in the community hygiene program. This suggests there may be benefits of more targeted community activities. However, this finding is tempered by the fact that for many indicators -- safe storage and use of drinking water, hand washing behaviors, perception about the causes of diarrhea, and perception about clean water – the positive improvement cannot be attributed to the project, even the community hygiene program. For a large majority of these indicators, the percentage of households exhibiting positive behaviors increased over the baseline, however both the participants and non-participants showed improvements so that the increase could not be attributed to the project.

A possible reason for this finding could be the short duration for which the community programs were implemented and the characteristics of the audience participating in the intensive hygiene activities. For example, programs such as the interactive theater and mother's training were designed for women with infants, but many participants were either old or very young women. However, this does not explain why both participants and non-participants exhibited improved behavior between baseline and endline.

A more plausible explanation could be that non-participants may have also benefited from the project and received behavior change messages through interactions with their friends, relatives and neighbors that participated in the program. Furthermore, the radio messages were relayed nationwide, which increases the likelihood that the respondents who did not hear the radio spots, benefited from acquaintances that did. Consequently, a comparison of outcome across the participants and non-participants would not find an impact. Another reason this may happen is the impact of other programs. Although, the impact of any TV spots aired by other programs were controlled for when estimating the impact, it is possible that other programs, for which we did not have data, could be positively impacting both participants and non-participants of the PSDW-HPP project.

References

1. Ahmad, T., and K. Alibhai. 2000. Health and Hygiene Education Programme: Northern Pakistan.
2. Ai. D, and E.C. Norton, "Interaction Terms in Logit and Probit Models," *Economics Letters* 80 (123-129), 2006
3. Asian Development Bank, "Water and Poverty: The Themes; A Collection of Thematic Papers", Prepared by partners in the Water and Poverty Initiative for the Third World Water Forum in Kyoto, Japan (March 16-23, 2003) and consultants for ADB under RETA 6093: Promoting Effective Water Management Policies and Practices, 62-75, January, 2004.
4. Donner, Allan, and Neil Klar, "Methods for Comparing Event Rates in Intervention Studies When the Unit of Allocation is a Cluster," *American Journal of Epidemiology*, 1994, Vol. 140(9).
5. Esrey, Steven, A., " Water, Waste and Well Being: A Multi Country Study" *American Journal of Epidemiology, 1999y* Vol. 143, No. 6: 608-623
6. Maluccio, John A. and Rafael Flores, "Impact Evaluation of Conditional Cash Transfer Program: The Nicaraguan *Red De Proteccion Social*," Food and Consumption and Nutrition Division, Discussion Paper No. 184. IFPRI, July 2004.
7. Nations Sanitation Policy (NSP), Government of the Islamic Republic of Pakistan, "National Sanitation Policy", Ministry of Environment, September, 2006.
8. Puhani, P.A., "The Treatment Effect, the Cross Difference, and the Interaction Term in Nonlinear "Difference-in-Differences" Models", IZA Discussion Paper No. 3478, 2008.
9. Ravallion, Martin, "The Mystery of Vanishing Benefits: An Introduction to Impact Evaluation," *World Bank Economic Review*, 2001, Vol. (15)1.
10. Skoufias, Emmanuel, "Is PROGRESA Working? Summary of the Results of an Evaluation by IFPRI," IFPRI Publication, September, 2001.
11. UNDP United Nations Development Programme and the Stockholm Environment Institute, Poverty-Environment Partnership, Linking Poverty Reduction and Water Management, 35-40, March, 2006a
12. UNDP, United Nations Development Programme "Human Development Report", Beyond Scarcity: Power, poverty and the global water crisis, 2006b Available at: http://hdr.undp.org.
13. UNICEF, "Basic Indicators," *The State of the World's Children 2006: Excluded and Invisible,* Table 1, Available at: http://www.unicef.org/sowc06/statistics/tables.php.
14. UNICEF, Joint Monitoring Program, http://www.wssinfo.org/en/122_definitions.html, 2008
15. WHO, "Combating Waterborne Diseases at the Household Level" 2007

Appendix A
Definition of Socio-Economic Class[17]

This appendix presents the details of a survey conducted by AC Nielsen in 2005 to establish a method of classifying rural and urban households into socio-economic classes (SECs).

A.1 Rural SEC Classification

Pakistan covers 797,000 square kilometers accommodating 130 million people out of which 87.8 million are living in 44,464 rural locations. To study and understand rural Pakistan a scale is needed; a socio-economic status scale which would help marketers and advertisers to more effectively target particular segments of the population with known profiles. It would also help research agencies to easily identify which socio-economic class an individual belongs to through a few simple questions which would not be too demanding on the skills of the interviewer or the knowledge of the respondent.

Background

Pakistan Advertisers Society (PAS) first initiated the idea of bringing together the marketers on one platform for market classification purposes in 1998. Resultantly a large scale Establishment, Media Habits and Socio-Economic Classification (SEC) for urban Pakistan was undertaken and published through AC Nielsen Pakistan. The SEC Grid developed through this survey has become a norm and is still used as the only currency across the country by the marketers as well as the researchers.

After the success of urban survey, PAS started receiving indications for a similar kind of survey in rural Pakistan which constitutes almost two-thirds of the country's population. The only source of information available for rural Pakistan was either population census 1998 or agriculture census1990. Both provide limited scope of information, as no data was available on durable ownership, FMCG consumption, agronomical practices, media habits or household profiling through any socio-economic classification of rural Pakistan.

On the other side, the world is becoming a global village and marketers need to understand markets in different countries, ESOMAR (European Society for Opinion and Marketing Research) is currently working toward harmonization of SEC scales throughout the world so that analysis of international research projects is more compatible and comparable.

With this purpose in mind, through the same platform of PAS and AC Nielsen, a similar kind of survey was conducted in rural areas of Pakistan in 2002. This reports led to the establishment of rural media habits via SEC survey, which are now available with the agency.

Socio-economic classification of rural Pakistan can help companies in:

- Determining current/potential market sizes for specific product categories
- Concentrating their efforts wherever a large percentage of the target market is located geographically

[17] Based on a survey conducted by AC Nielsen in Pakistan under the aegis of the Pakistan Advertisers Society.

- Developing better advertising copy, tailor-made to the profile of the individuals in various socio-economic classes
- Developing more cost effective and efficient media plans.

Objective

The objective of conducting a socio-economic classification exercise in rural Pakistan was to establish the rural household profile, to understand media habits of 12+ year rural individuals and to segment the rural populace on the basis of the most relevant/distinguished variables for the purpose of target marketing

In addition to the above, a screening criterion was also to be established whereby the field interviewers can easily and quickly determine which rural individual/household belongs to which socio-economic class.

Research Design

The study was conducted in two steps through face-to-face interviews with the selected households of the randomly identified rural areas. First, household level data (i.e., penetration of various durables, demographic data, and other household features) were obtained through "The Establishment Survey". As a second step, information about media habits at individual levels was sought via "The Media Habits Survey".

To obtain household level data, our target respondent was the male head of a selected household or the chief wage earner of that household. To ascertain individual media habits, our target respondents were adult males and females over the age of 12.

For this survey, a probability based, multi-stage, disproportionate, systematic random sample design was adopted by using Probability Proportionate to Size (PPS) technique.

The universe constituted all non-institutional households in rural Pakistan. The primary sampling units (PSUs) were villages. The secondary sampling units (SSUs) were households and individuals. A household may consist of a single person living alone or a group of persons who normally live and eat together. A total of 5,998 households were selected in 500 villages across Rural Pakistan. Two interviews were conducted in each target household, one with head of household and another with a randomly selected 12+ year male or female member of that particular household. The data was analyzed using cluster and factor analysis to reach a data based classification system for urban households in Pakistan. The following pages describe the results.

COMPOSITION OF SEC GRID:

Socio-economic classes have been identified using two variables Education of Head of Household and Type of House.
Head of household is the one who is the decision-maker in majority of household matters like marriage, purchase and sale of property etc. In rural Pakistan, structure of house is also a solid indicator about the socio-economic status of the household.

Previously stated monthly household or personal income was being used as a classification method in both urban and rural Pakistan. Some companies also used to gather information on the availability of a long list of household durables and were calculating indices to segment the target population.

However, income is not considered a suitable indicator for classification purposes. The main reasons are that income may be overstated or understated either intentionally or unintentionally and it is generally not stable over time. Another problem is non-response, as people prefer not to disclose their income for different reasons. On the other side, in contrast to urban areas where monthly household income can be stated upfront, it is almost impossible for an agriculturist household to state monthly income due to the five or six monthly nature of crops. Therefore, in case of rural households, it is impractical to use income as a suitable indicator for classification purposes.

Secondly, asking for ownership of durables, though a stable indicator, involves long lists making the job tedious for surveyors and then for analysts in calculating the indices and segmenting populations.

After various analyses, the rural SEC Grid emerged as a two-dimensional matrix constructed with the following two variables:

Levels of head of household's education
Type of house

The four types of house are taken on the horizontal axis whereas seven levels of education are taken on the vertical axis, thereby forming a 28 cell grid.

The four types of house are given below. These are to serve as broad guidelines for the ease of field interviews.

Type of House:

Kuchha: House where both the roof and walls are made of Kuchha material

Semi Pukka: House where EITHER the roof or the walls are made of Pukka material

Pukka Lower: House where both the roof and walls are made of Pukka material but EITHER the Kitchen, Toilet or both are not present.

Pukka Upper: House where both the roof and walls are made of Pukka material, and both Kitchen and Toilet are present.

For classifying any rural household among the above mentioned types of house, the surveyors would need to ask about the material used for construction of walls and roof of the rooms. The material which categorizes pukka and kuchha is mentioned below:

Table A-1: House Structure and Materials Used

Structure	Material used for construction
***Pukka* Walls**	Burnt Bricks, Plastered Walls, Bricks-Mud, Wood-Stone, Galvanised Iron or other Metal Sheets, Concrete
***Kuchha* Walls**	Mud, Wooden, Grass, Leaves, Reeds,Wooden/Grass Fence
***Pukka* Roof**	Concrete Slab (RCC), Iron Girder, Wooden Beams/Bricks/Stones, Iron Sheet, Mud and Stones, Prefabricated Slab
***Kuchha* Roof**	Wood, Mud, Unburnt Bricks-Mud, Wood/Bamboo, Reeds, Grass, Leaves, Straws

The seven levels of education and what each stands for is discussed below:

Education:

Illiterate: An individual who cannot read or write even simple/plain Urdu, and has never attended any school.

Up to primary: An individual who has not attended formal school but can read simple Urdu or has less than 5 years of formal schooling.

School 6-9 years: An individual who has attended and cleared 5 years of primary school, has obtained regular/formal religious education or has attended school for 6, 7, 8 or 9 years.

Matric: An individual who has passed/cleared matriculation / Ordinary Level GCSE examination and therefore has attended formal school for 10/11 years.

Intermediate: An individual who has attended school/college for 12 years and has passed/cleared intermediate (F.A., F.Sc., I.Com.) / Advanced Level GCSE examination.

Graduate: An individual who has attended school / college for 14 years and holds a Bachelor of Arts (B.A.) / Sciences (B.Sc.) / Commerce (B.Com.) / Business Administration (B.B.A.) / Law (LL.B.) degree.

Post-Graduate: An individual who has done a Masters in Arts (M.A.), Masters in Science (M.Sc.), Masters in Commerce (M.Com.), Masters in Computer Sciences (MCS), Masters in Business Administration (MBA), MBBS, Chartered Accountancy (CA), Bachelor of Engineering (BE), FRCS, MRCP, Ph.D. or any other higher studies within the country or abroad.

DESCRIPTION OF SOCIO-ECONOMIC CLASSES

In order to provide marketers and researchers with more tangible/concrete insight into the types of individuals to be encountered in each socio-economic class a brief description of each category has been compiled below:

A This is the most educated class in rural Pakistan where the education of the head of household is at least intermediate and the structure of house is either pukka lower or pukka upper.

B A high percentage of individuals in this class have acquired education upto matriculation level and the structure of house is any one from all four types.

C This is the middle class of rural Pakistan. Education level of heads of households is much lower than in SECs A and B.

D This is the largest SEC in terms of number of households. Illiteracy among the heads of households is very common. Structure of house is either semi pukka or pukka lower.

E Most of the heads of households in this class have not acquired any formal education. Structure of the house is kuchha.

USAGE OF SEC GRID

The seven questions, which are to be enquired from any survey respondent, are given below.

Q.1 who is the head of your household? By head of household I mean, the person who is responsible for the major decisions related to the household matters, family, etc.?

Respondent himself/herself……[1]

Any other (note down)_____

Q.2 What is the highest level of education that you / your head of household has attained? (Single Code)

Illiterate	[1]
Up to primary (1-5 classes)	[2]
6-9 classes pass	[3]
Matric	[4]
Intermediate/I.Com/D.Com/F.A/F.Sc.	[5]
Graduate	[6]
Post Graduate	[7]

Q.3 From what material is the roof of most of the rooms of your household made?

Pukki		Kachhi	
Concrete slab (RCC)	[1]	Mud	[7]
Iron girder, wooden beams/bricks, stones	[2]	Wood-mud	[8]
Iron sheets	[3]	Wood/bamboo	[9]
Stones, mud	[4]	Grass/leaves, straws	[10]
Prefabricated Slab	[5]		
Girder, TR and bricks	[6]		

Q.4 Now please tell me what is the type of most of the walls of your household?

Pukki		Kachhi	
Bricks and cement	[1]	Mud	[6]
Bricks	[2]	Wooden	[7]
Wood-stone, mud	3]	Grass, leaves/grass fence	[8]
Iron or other metal sheets	[4]		
Concrete	[5]		

Check from Q.3 and Q.4, if both roof and walls of the most of the rooms are made of pukka material then ask next questions otherwise go to Q.7

Q.5 Is there a proper kitchen in your house i.e., a separate room which is only used for cooking purpose?

Yes .....[1]

No.....[2]

Q.6 Do you have a toilet in your household?

Yes .....[1]

No.....[2]

Q.7 (Surveyors record the type of house in the grid given below after looking at Q.3-Q.6)

Type of House	Code	Definition
Kuchha	1	Both roof and walls are made of kuchha material
Semi Pukka	2	Either the roof or walls are made of Pukka material
Pukka Lower	3	Both roof and walls are made of pukka material but either the kitchen, toilet or both are not present
Pukka Upper	4	Both roof and walls are made of pukka material and both kitchen and toilet are present

All interviewers must have the SEC Grid, list of structure of house and education codes with them before they go into the field. Without these, the interviewers will not be able to determine the SEC.

Moreover, the respondent must not be exposed to these materials. Both the questions must be discreetly asked. No cards must be shown to the respondent, as he/she is not expected to understand what kinds of houses are classified under say pukka lower category or what educational qualifications fall under 'post-graduate' category.

Don't know

If any respondent does not know what the education level of the head of household is or from what material roof of the rooms and walls of the household are made then such a respondent/household should be dropped from the sample.

The two variables which determine the SEC of any rural household i.e., Education and structure of house, are fairly stable demographic variables. The SEC system developed for rural Pakistan is based on these variables and it is a stable system. At present many large multinationals such as Pakistan Tobacco Company, Reckitt & Colman, Procter & Gamble, etc. are using these SEC definitions in their market research.

All companies and research organizations can use these socio-economic classes to design samples and select individual respondents.

A.2 Urban SEC Classification

Introduction

Pakistan covers 797,000 square kilometers accommodating 130 million people out of which 42.4 million are living in 466 urban locations. To study and understand a diverse populace such as that of urban Pakistan a scale is needed; a socio-economic status scale whereby marketers and of course marketing researchers can easily select the segments most appropriate for their products and services.

The motivation for a uniform socio-economic classification system is derived from the need for a standard criterion for basic segmentation of markets to ensure comparability across studies from different sources and at different times. The manufacturer who commissions a U&A study that tells him that his best opportunities are with classes 'A' and 'B', for instance, will certainly want to be sure that the group discussions he carried out or intends to carry out and the media consumption data he bought from different sources all refer to the same socio-economic classes.

Background

Segmentation of any market can be done based on one or more variables, which may be demographic, social, psychographic, or economic in nature. To date, stated income is being used for segmenting the consumers in Pakistan for research as well as marketing purposes. There are, however, numerous problems in using this criterion. First, income is typically understated or overstated by respondents either intentionally or unintentionally. Second, if the respondent is a housewife many times she does not know the correct household income. The errors ensuing because of such misreporting cannot be rectified. No exact factor can be assigned to misreporting. Other issues are related to non-response, and obsolescence of the income data for tracking studies. Hence, both researchers and research users in Pakistan had felt the need for reviewing the situation and coming up with some relevant, reliable, consistent and practical criteria, which will overcome the pitfalls of income criteria.

Pakistan is not alone in facing these difficulties. Because of similar limitations, most countries of the world have moved to using a basic socio-economic classification of households for basic market segmentation. Different variables are used to define these socio-economic classes. The most commonly used variables are; occupation and/or education of the chief earner, type/number of durables present at home, type/nature/ownership of place of dwelling, number of bathrooms in the house, number of domestic servants, education of the housewife, etc.

Since the world is becoming a global village and marketers need to understand markets in different countries, ESOMAR (European Society for Opinion and Marketing Research) is currently working toward harmonization of SEC scales throughout the world so that analysis of international research projects is more compatible and comparable.

With this purpose in mind, Pakistan Advertisers Society (PAS) decided to conduct a socio-economic classification exercise for urban Pakistan. Socio-economic classification can help companies in:

Determining current/potential market sizes for specific product categories

Concentrating their efforts wherever a large percentage of the target market is located geographically

Developing better advertising copy, tailor-made to the profile of the individuals in various socio-economic classes

Developing more cost effective and efficient media plans.

Objective

The aim of conducting a socio-economic classification exercise was to first identify distinct socio-economic classes as they exist in urban Pakistan, second to determine the size of each class, and third to determine key characteristics of each class.

In addition to identifying, sizing and profiling the classes, a screening criterion was also to be established whereby the field interviewers can easily and quickly determine which individual/household belongs to which socio-economic class.

Research Design

To meet the objective first a pilot study and then a large quantitative, face-to-face survey was conducted. This survey was divided into two steps. In each sampled household two interviews were conducted, one with the housewife of the household and one with a randomly selected 12+ years male/female household member.

Probability-based, multi-stage, stratified, disproportionate sample design was used for this survey.

The primary sampling unit was clusters, which are systematic intervals in the electoral rolls as defined by the administrative structure of local Councils/Election Commission. The secondary sampling unit were households, consisting of a single person living alone or a group of persons who normally live and eat together. In this survey a total of 40,228 interviews were conducted in 20,114 households across 116 cities representing all the four provinces. The data were analyzed using cluster and factor analysis to reach a data-based classification system for urban households in Pakistan. The following pages describe the results.

Composition of SEC Grid

Socio-economic classes have been identified using the social variable of the education of the chief earner and the economic variable of the occupation of chief earner. The chief earner is the member of a household who contributes the most to the budget of the household and bears the greatest proportion of the overall household expenses. The rationale behind classifying a household on these two variables is that the consumption, income and lifestyle of a household is strongly correlated with how educated and economically sound the chief earner is. Moreover, in many households the chief earner is also the head of the household. In such cases, the behavior/opinion/attitudes of the chief earner may influence the behavior and lifestyle of all other household members.

The SEC Grid is a two-dimensional matrix constructed with the following 2 variables:

Levels of chief earner's education

Occupation categories of chief earner

The seven levels of education are taken on the horizontal axis and the eleven occupation categories are taken on the vertical axis, thereby forming a 77 cell grid. The seven levels of education and what each stands for is discussed below:

Education

Illiterate: An individual who cannot read or write even simple/plain Urdu, and has never attended any school.

Less than Primary: An individual who has not attended formal school but can read simple Urdu, or has less than 5 years of formal schooling.

School 5-9 years: An individual who has attended and cleared 5 years of primary school, or has obtained regular/formal religious education, or has attended school for 6/7/8/9 years in classes above the primary level.

Matric: An individual who has passed/cleared matriculation/Ordinary Level/GCSE examination and therefore has attended formal school for 10/11 years.

Intermediate: An individual, who has attended school for 12 years, is Fellow of Arts/Sciences/Commerce (F.A., F.Sc., I.Com.).

Graduate: An individual who has attended school for 14 years, i.e., is a Bachelor of Arts (B.A.)/Sciences (B.Sc.)/Commerce (B.Com.)/ Business Administration (B.B.A.)/Law (LL.B.)

Post-Graduate: An individual who has done Masters in Arts (M.A.), Masters in Science (M.Sc.), Masters in Commerce (M.Com.), Masters in Ph.D. or has done any other higher studies within the country or abroad.

The 11 broad occupation classifications and details of the kind of workers, employees and self-employed people falling in each are given in the following page. These are to serve as broad guidelines for the ease of field interviews. The occupation categories are listed such that occupational status is in line with the socio-economic status. That is households where the chief earner is an unskilled worker tend to belong to the lower stratum of the society, whereas households where the chief earner is a large businessman will belong to the upper-most stratum of the society.

Occupation Classification

Unskilled Worker: Workers who largely do not handle machinery or sophisticated instruments and do not require special training or diplomas, e.g., manual labourer, peon, doorman, fisherman, waiter, domestic servant, ward boy, messenger, helpers in shops, other establishments, loaders, cook, newsboy, agricultural labourer.

Petty Trader: Traders and persons engaged in selling petty items or personal services without having any properly constructed ('pukka') shop/establishment, e.g., hawkers, street vendors, 'pan'/cold drink shop owners, peddlers, tea/coffee/juice stall owners, etc.

Skilled Worker: Workers who handle machinery or require special training/diplomas, e.g., carpenters, chefs, electricians, drivers, mechanic, technician, tailors, armed guard, repairmen, telephone operators, computer operator, cobbler, barber, farmer, steward, typist, overseas workers, craftsman, nurse, LHV, dispenser, moazzan, lowest designations in police/armed forces, e.g., jawans, hawaldar, sipahi, batman.

Non-Executive Staff: The category includes white-collar workers, such as clerks, salesmen.

Supervisory Level: Those in supervisory/regulatory positions who are not senior enough to be called officer/executive, e.g., head constable, station master, shop managers, primary school teachers, imam masjid/preacher, supervisors working in factories/offices.

Small Shopkeeper/Businessmen: This category encompasses people engaged in providing retail, restaurant or personal services and operate from a properly constructed (i.e., 'pukka') establishment, e.g., general/'kiryana' store owners, general merchants, butcher, all small shop owners (e.g., laundry shops, cloth merchants, shoe shops, hair dresser/beauty parlor) real estate agents, small hotel owner (e.g., small tea shops, tandoors).

Lower/Middle Executive/Officer: Employees of grade 14-17, high school/college teachers and lower managerial positions in private companies.

Self-Employed/Employed Professionals: Accountants, doctors, engineers, lawyers, architects, actors, brokers, editors, journalists, trainers, authors, university teachers/professors, players and all other professionals who are either employed or have their own private practices. 'Hakeems', homeopathic doctors are also included in this category.

Medium Businessmen: Owners of small companies/big shops/departmental stores, jewelers, car showroom owners, owners of air-conditioned hotels/restaurants, small-scale factory owners.

Senior Executive/Officers: MDs, CEOs, directors, senior government officials, employees of grade 18 and above, DG's secretaries, etc. Overseas Pakistanis working as executives, managerial positions in private companies.

Large Businessmen/Factory Owner: Landlord, industrialist, big contractors/importers/exporters, owners of factories, owners of chain of hotels/restaurants.

Description of Socio-Economic Classes

In order to provide marketers and researchers with more tangible/concrete insight into the types of individuals to be encountered in each socio-economic class a brief description of each category has been compiled.

SEC Households in which the chief earners are:

A_1 Well educated, self-employed/employed professionals, senior level executives/officers in public/private limited organizations, well-educated medium-to-large scale businessmen.

A_2 Relatively less well educated, medium-to-large scale businessmen and professionals. Well educated middle level executives, small businessmen and supervisors.

B Relatively less well-educated lower/middle level executives and officers, well educated small businessmen and supervisors.

C Predominantly small retailers/businessmen, supervisors and lower level executives who have 5-10 years of schooling.

D Relatively well educated skilled workers; not so well-educated small retailers and non-executive staff members.

E_1 Skilled/unskilled workers, petty traders and non-executive staff members who have at least 5-10 years of schooling.

E_2 Predominantly, illiterate unskilled/skilled workers and petty traders.

Usage of SEC Grid

The SEC Grid can easily be used by all fieldworkers. The two questions, which are to be enquired from any survey respondent, are given below.

Q1. Could you kindly tell me what is the occupation of the chief earner of your household, i.e., what is his/her job designation/nature of business? By chief earner, I mean the member of the household who bears the greatest percentage of the overall household expenditure.

PROBE THE RESPONDENT FOR ALL DETAILS AND NOTE THE VERBATIM.

INTERVIEWER: CODE WITH THE HELP OF 11 OCCUPATION CATEGORIES GIVEN OR PLEASE REFER TO THE NOTE.

Unskilled Worker	[1]	
Petty Trader	[2]	
Skilled Worker	[3]	
Non-Executive Staff	[4]	
Supervisor	[5]	
Small Shopkeeper/Businessman	[6]	
Lower/Middle Officer/Executive	[7]	
Professional (self employed or in service)	[8]	
Medium Businessman	[9]	
Senior Executive/Officer	[10]	
Large Businessman/Factory Owner	[11]	
Retired	[97]	→[GO TO NOTE 1]
Don't know	[98]	→[TERMINATE]
Student/housewife/unemployed	[99]	→[GO TO NOTE 3]

Q2. What is the educational qualification of the chief earner of your household?

Illiterate	[1]	
Less than Primary	[2]	
School 5-9 years	[3]	
Matric	[4]	
Intermediate	[5]	
Graduate	[6]	
Post-graduate	[7]	
Don't know	[9]	→[TERMINATE THE INTERVIEW]

All interviewers must have the SEC Grid, list of occupation and education codes with them along with the note before they go into the field. Without these, the interviewers will not be able to determine the SEC.

Moreover, the respondent must not be exposed to these materials. Both the questions must be discreetly asked. No cards must be shown to the respondent, as he/she is not expected to understand what kinds of jobs are classified under say non-executive staff category or what educational qualifications fall under 'post-graduate' category.

If the chief earner in a house has done his matriculation and is a doorman, according to the grid, that household will be classified as an E_1 class household.

Similarly, if the chief earner in a household has done an MBA and is working in a private firm as an executive, this household will be classified as a household belonging to the A_2 Class.

In this manner, each sample household/individual can be classified into one or another socio-economic class.

In case the reply of the respondent is other than the pre-coded options given then the field workers must refer to the instructions in the following note.

Note:
If the respondent states that the chief earner of his/her household is; retired, a student, housewife, unemployed or he/she does not know what the chief earner's occupation is or he/she says that income comes in from rent or land, please refer to the following instructions:

Retired:
If the respondent says that the chief earner is retired then ask the respondent what was the chief earner's occupational status immediately before retirement. After asking about the chief earner's education, using the SEC Grid determine the socio-economic class of that household.

Don't know:
If any respondent does not know what the occupation of the chief earner (C.E) is or from what source income is earned to meet household expenses and or the education of the C.E then such a respondent/household should be dropped from the sample.

Student/Housewife/Unemployed:
If the C.E of a household happens to be a student, housewife or an unemployed person then ask the survey respondent as to how the expenditures (food, clothing, rent, etc.) are being met or what their source of income is;

(a) If the source of income is rent on any kind of property and or interest on a bank account then ask the respondent as to who was the chief earner of the household. That is who left the property or money in the bank account for the descendants who are using the money. Then ask about the education and occupational status of that individual and code the appropriate SEC accordingly.

(b) If the source of income is irrigated/agricultural land then ask the respondent about the total land area the household owns. If the respondent does not know this upfront, he/she may ask another available member of the household.

(i) If the total land area is less than 12 and a half acres then the occupation category of the chief earner will be agricultural worker or "unskilled worker".

(ii) If the total land area is, 12 and a half acres or more but less than 50 acres then the occupation category of the chief earner will be farmer or "skilled worker".

If the total land area is 50 acres or more then the occupation category of the chief earner will be "large businessmen".

Occupation and education are fairly stable demographic variables. As the SEC system developed for urban Pakistan is based on these variables, it is a relatively stable system. At present many large multinationals such as Pakistan Tobacco Company, Reckitt & Colman, Procter & Gamble, Coca-Cola Export Corporation, Rafhan, etc. are using these SEC definitions in their market research.

Results

Urban Socio-Economic Classification (SEC)

Urban households are classified into different SEC based on the education and occupation of Chief wage earner.

Chief wage earner is the person who contributes the most to the household budget.

Table A-2: SEC Class by Occupation of Chief Earner

Occupation of Chief Earner	Education of chief earner						
	Illiterate	Less than Primary	School 5-9 years	Matric	Intermediate	Graduate	Post Graduate
Unskilled worker	E-2	E-2	E-1	E-1	D	D	C
Petty traders	E-2	E-2	E-1	E-1	D	C	C
Skilled workers	E-2	E-2	E-1	D	D	C	C
Non-executive staff	E-2	E-2	D	D	D	C	C
Supervisory level	D	D	C	C	B	B	B
Small shopkeeper/Businessmen	D	D	C	C	B	B	A-2
Lower/Middle: Executive, Officer	D	C	C	C	B	B	A-2
Self employed/Employed/Professionals	B	B	A-2	A-2	A-2	A-1	A-1
Medium Businessmen	B	A-2	A-2	A-2	A-2	A-1	A-1
Senior Executive/ Officer	B	A-2	A-2	A-2	A-1	A-1	A-1
Large Businessmen/Factory owner	A-2	A-2	A-2	A-1	A-1	A-1	A-1

Source: Based on the survey conducted by AC Nielsen Pakistan for PAS in 1998.

Description of Socio-Economic Classes

A1:

Well-educated, self-employed/employed professionals, senior level executive/officers in public/private limited organizations, well-educated medium to large-scale businessmen.

A2:

Relatively less well educated, medium to large scale businessmen and professionals. Well educated middle level executives, small businessmen and supervisors.

B:
Relatively less well-educated lower/middle level executives and officers, well-educated small businessmen and supervisors.

C:
Predominantly small retailers/businessmen, supervisors and lower level executives who have 5-10 years of schooling.

D:
Relatively well educated skilled workers; not so well educated small retailers and non-executive staff members.

E1:
Skilled/unskilled workers, petty traders and non-executive staff members who have at least 5-10 year of schooling.

E2:
Predominantly, illiterate unskilled/skilled workers and petty traders.

Table A-3: Share of SEC in Urban Households

SEC	percent in Urban Pakistan (1998)	Avg. Household Income Rs. (1998)	*Avg. Income $ (1998)	percent in Urban Pakistan (2005)	Avg. Household Income Rs. (2005)	*Avg. Income $ (2005)
A1	2.8	16,561	368	4.3	25,217	422
A2	3.8	10,134	225	5.5	17,485	292
B	10.0	9,418	210	12.4	12,475	209
C	18.5	6,873	153	20.0	9,508	159
D	21.6	5,789	129	22.4	8,019	134
E1	19.4	4,385	98	15.0	6,490	109
E2	23.9	4,007	89	20.4	5,779	97

Source: Based on survey conducted by AC Nielsen Pakistan for PAS in 1998 and 2005.

Table A-4: Urban Households by SEC by Province (2005)

Classes	Punjab	Sindh	KPK	Baluchistan	Total
A1	166,349	103,143	22,567	14,503	**306,562**
A2	222,661	116,604	37,319	14,538	**391,123**
B	461,777	340,603	58,907	28,813	**890,099**
C	854,977	457,665	76,115	40,435	**1,429,191**
D	895,853	562,455	92,547	53,421	**1,604,277**
E1	650,610	339,153	58,461	24,228	**1,072,451**
E2	765,741	527,556	98,924	65,630	**1,457,851**
Total	**4,017,967**	**2,447,178**	**444,840**	**241,569**	**7,151,553**

Source: Based on survey conducted by AC Nielsen Pakistan in 2005.

Rural Socio-Economic Classification

Rural SEC is based on two socio-economic variables: Education of the Head of household and Structure of the house.

Table A-5: Rural SEC Classification

Education	Structure of House			
	Kuchha	Semi Pukka	Pukka lower	Pukka Upper
Illiterate	E	D	D	C
Up to Primary	E	D	C	C
School 6-9 years	D	C	C	B
Matric	D	C	B	B
Intermediate	C	C	B	A
Graduate	C	C	A	A
Post Graduate	B	B	A	A

Source: Household numbers are based on Pakistan 1998 Census Report of Pakistan. EC shares are based on the survey conducted by AC Nielsen Pakistan for PAS in 2002.

Definitions of Type of Houses

Kuchha:
House where both the roof and walls are made of kuchha material.

Semi Pukka:
House where EITHER the roof or the walls is made of pukka material.

Pukka Lower:
House where both the roof and walls are made of pukka material, but the kitchen, toilet, or both are not present.

Pukka Upper:
House where both the roof and walls are made of pukka material, and both kitchen and toilet are present.

Table A-6: Share of SEC in Rural Households

SEC	percent in Rural Pakistan	Household	Avg. Income Rs.	*Avg. Income $
A	4.2	533,362	9,841	165
B	10.3	1,328,049	6,629	111
C	24.9	3,198,027	5,417	91
D	31.6	4,054,833	3,801	64
E	29.1	3,733,531	3,321	56

Source: Based on Survey conducted by AC Nielsen Pakistan for PAS 2002.

Table A-7: Rural Households by SEC and Province

Classes	Punjab	Sindh	KPK	Baluchistan	Total
A	255,537	190,066	78,994	8,550	533,362
B	931,960	166,575	211,363	19,239	1,328,049
C	2,061,477	600,096	448,346	87,642	3,198,027
D	2,647,710	600,096	638,359	168,872	4,054,833
E	1,481,687	1,294,157	465,426	491,651	3,733,531
Total	7,378,371	2,850,990	1,842,488	775,954	12,847,802

Source: Household numbers are based on Pakistan 1998 Census Report of Pakistan. SEC shares are based on the survey conducted by AC Nielsen Pakistan for PAS in 2002.

Appendix B
Sampling Methodology and Data Collection

This appendix provides the details of the sampling methodology and the sampling performance, particularly in high risk areas.

Sampling Frame for the Survey

As per the initial phase of the program, there were total 31 treatment districts and within these 31 districts a total of 126 UCs. Therefore, the treatment sample was drawn from these districts. If one uses an overlap of approximately 70 percent when conducting the endline (post-program) survey, a relatively conservative assumption on the correlation across the control and treatment sample in the absence of information on it of 0.4, the total number of observations that would detect a 5 percent difference in percentage points for key indicators is 1237. To allow for response rate, a sample size of 2000 was targeted for each of the control and treatment areas.

Sample Selection – Treatment Group

Provinces were treated as a stratum in this design. This allowed independent selection of the sample in each province. Then the total sample for UCs was selected in proportion to the population of the stratum. This was implemented by sorting the UCs by district and *tehsil*. Next, a random number between 0 and 1 was generated which was multiplied by the sampling interval. The sampling interval was obtained by taking the total population size in the province and divided by the number of UCs to be selected. To this number, the sampling interval was added successively until the target number of UCs was obtained. Each random number was associated with the cumulative total number associated with that UC. If the random number was smaller than the cumulative total associated with the UC but was larger than the previous cumulative total, then that UC was selected.

In each UC that had a probability of selection less than one, 28 households were selected; and for UCs where the probability of the selection was greater than one, a maximum of 60 households were selected.

Sample Selection – Control Group

To serve as control for the treatment districts, control districts were selected that were most similar to each other using the five key indicators that were expected to affect program outcomes and for which secondary data was available. In rough order of priority the five key indicators are:

1) Adult female literacy
2) Diarrhea outcomes
3) Access to tap water
4) Presence of flush toilets
5) Enrollment in primary government schools

Some control districts were naturally a better match than others, and these were graded based on how closely they matched with the treatment district on the abovementioned variables. When picking *tehsils* from control districts, *tehsils* were chosen that were geographically the most proximate, although not immediately contiguous, to avoid contamination. Once the *tehsil* was chosen, the UC council that was the *tehsil* headquarters was chosen since all treatment UCs are by design *tehsil* headquarters where the

water filtration systems have been built. For each treatment UC/*tehsil* in the sample, at least three possible *tehsils* were picked from the best- matched control districts using information on rural/urban UC.

Geographical Coverage and Sample Size

A total of 4,000 interviews were conducted, spread throughout the four provinces and AJK and FATA region with an equal split between the test and comparison group.

The sample split for the test and comparison group is provided in Tables A-1 and A-2 below. Districts highlighted in red are those that could not be sampled because of high risk conditions.

Table B-1: Union Council Sample for Treatment Group

Province	District	*Tehsil*	Union Council	Sample
KPK	Batagram District	Allai	2.Alai	28
	Buner District	Daggar *Tehsil*	Dagar	28
	Kohistan District	Pallas	3. Pallas	28
	Mansehra District	Bala kot *Tehsil*	2. Bale(Balakot)	28
		Mansehra *Tehsil*	3. Mansehra	43
		Oghi *Tehsil*	1.Oghi	28
	Shangla District	Alpuri tehs	3.Alpuri	28
		Chakisar *Tehsil*	1.Chakesar	28
	Swat District	Matta *Tehsil*	2.UC Matta -02	28
	Upper dir District	1.urban dir	1.Urban Dir	28
		2.wari	2.Main Wari	28
BALUCHISTAN	Jaffarabad District	Usta Muhammad	4.UC 1 Usta Muhammad	28
		Subatpur	5.UC Subatpur	28
	Kech District	Turbat	6/A.UC Upsar Turbat City 6/B.UC Kolwahi Bazar	28
		Zamarun	5.UC Zamarun	28
	Lasbela District	Dureji	5. UC Dureji	28
		Hub	3. UC Hub	55
		Uttal	2/A. UC Uttal 2/B. UC Liari Bazar	28
	Zhob district	Zhob Sub-Division	UC Zhob	28
PUNJAB	Rawalpindi District	Gujar Khan	55/1 Urban	28
		Kahuta	Shumali Urban	28
		Kotli Satian	UC 70	28
		Taxila	UC55/2	28
		.Kallar Sayyedan	47 Choha khalsa	28
	Okara District	.Depalpur	UC Depalpur	28
		Renala Khurd	UC1&2	28
	Khushab District	Khushab	25 Johar Abad	28
	Khanewal District	.Jahanian	Urban 26	28
		Kabirwala	Kabirwala City	28
	Dera ghazi khan District	Tribal Area	Tribal Area	118
		Tunsa Sharif	*Tehsil* 13 H/Q	28
	Gujrat District	Gujrat	UC56	28
		Kharian	Kharian	28
	Lahore District	Wagha	Uc# 12 Naseer abad	28
		Nishtar	Uc#135 Ismail Nagar	28
A.J.K.	Bagh District	Bagh	UC Bagh	28
	Muzaffarabad District	Garhi Dupatta	Hatian Dupatta	28
		Hatian Bala	Mera Kala	65
		Muzaffarabad	Domail Spring	65
	Neelam District	Sharda	UC Sharda	28
	Poonch District	Rawalakot	CMH Rawlakot	28
		Hajera	UC Hajera	28
		Abbass pur	Darra	30
SINDH	Jamshoro District	Manjhand	Manjhand	28
		Kotri	Kotri UC1	28
		Sehwan Sharif	Sehwan Sharif 1	28
	Dadu District	Dadu	Dadu UC Two	28
		Johi	Johi UC One	28

Table B-1: Union Council Sample for Treatment Group

		K.N. Shah	K.N. Shah	28
		Mehar	Mehar UC	28
	Khairpur District	Gambat	Gambat UC2	28
		Kingri	Pir Goth	28
		Kot Digi	Kot Digi	28
		Nara	Nara UC Imam Bargah	28
	Larkana District	Dokri Taluka	Badah	28
		Larkana Taluka	UC 5	28
		Ratodero Taluka	Rato dero	28
	Sukkur District	Rohri Taluka	Rohri UC21	28
		Salehpat Taluka	Saleh Pat	28
		Old Sukkur Taluka	Sukkur UC4	28
	Thatta District	Jati Taluka	Jatti UC1	28
		Keti Bunder Taluka	Ketti Bandar	28
		Mirpur Bathoro Taluka	Mirpur Bathoro	28
		Shah Bunder Taluka	Chuhar Jamali	28
		Sujawal Taluka	Sajawal UC1	28

Table B-2: Union Council Sample for Control Group

Province	District	*Tehsil*	Union Council	Sample
KPK	Mardan	Takht bai	Makori	28
	Malakand Protected Area	Sam rani zai sub-division		28
	Charsadda District	Tangi	Uc tangi	28
	Swabi District	Lahor *tehsil*		28
		Swabi *tehsil*		43
	Abbottabad district	Abbottabad *tehsil*		28
	Chitral District	Chitral sub-division		28
		Mastuj sub-division		28
	Haripur District	Ghazi	Ghazi	28
	Lower Dir District	Temergara sub-division		28
		Jandool sub-division		28
BALUCHISTAN	Bolan District	Bhag	Jalal Khan	28
		Dhadar	Dhadar	28
	Gwadar District	Gwadar Sub-Division		28
		Pasni Sub-Division		28
	Khuzdar District	Wadh	2 Gazgi	28
		Khuzdar	1 Ferozabad	55
	Kharan District	Kharan sub-Division		28
	Musakhel District	Musakhel subdivision		28
AJK	Sudhnati district	Pallandari	Pallandari	28
	Kotli district	Fatehpur	Uc fateh pur	28
		Sehnsa	Uc sehnsa	65
		Kotli	Uc kotli	65
	Bhimber district	Samahni	Uc samahni	28
	Mirpur district	Mirpur	Uc 2	28
		Dudyal	Uc dudyal	30
PUNJAB	Chakwal District	Chakwal *Tehsil*		28
		Choa saidan shah		28
		Tala gang *Tehsil*		28
	Jhelum District	Jhelum *Tehsil*		28
		Pind dadan khan		28
	Kasur District	Kasur	Kasur No. 1	28
		Pattoki	Pattoki No. 2	28
	Mianwali District	Mianwali *Tehsil*		28
	Layyah District	Karor lal esan *Tehsil*		28
		Layyah *Tehsil*		28
	Muzaffargarh District	Muzaffargarh	Muzaffargarh City - 4	118
		Jatoi	Shaher Sultan	28
	Mandi Bahauddin District	Malakwal	UC No. 49	28
		Mandi Bahauddin	UC No. 5	28
	Gujranwala District	Gujranwala city *Tehsil*		28
		Wazirabad *Tehsil*		28
SINDH	Mirpur khas District	Digri taluka		28
		Kot Ghulam Mohammad taluka		28
		Mirpur khas taluka		28
	Nawabshah District	Nawab shah	Nawab shah 1	28
		Daulat pur	kazi Ahmed 1	28
		Sakrand	Sakrand 1	28
		Daur	Jam Sahib	28

Table B-2: Union Council Sample for Control Group

Province	District	*Tehsil*	Union Council	Sample
	Jacobabad District	Garhi khairo taluka		28
		Jacobabad taluka		28
		Kandhkot taluka		28
		Kashmore taluka		28
	Shikarpur Ddstrict	Shikarpur	M C Shikarpur - I	28
		Khanpur	U C Khanpur	28
		Lakhi	U C Lakhi	28
	Ghotki District	Daharki	Daharki (Urban)	28
		Ubauro	Ubauro (Urban)	28
		Ghotki	Ghotki I (Urban)	28
	Badin District	Badin	Badin I	28
		Talhar	Talhar	28
		Matli	Matli I	28
		Tando bago	Tando Bago	28
		Golarchi	Golarchi	28

Sampling Performance

A summary is given below in table A-3 below, which provides a detailed breakdown among both treatment and control groups. In the KPK provinces two districts that were under a civil war situation saw the situation worsen during the survey period, and could not be covered. Those districts were Shangla and Swat in KPK, and accounted for 84 households (4 percent of the sample).

Table B-3: Achieved versus Planned Sample

Province	Treatment		Control	
	Planned	Achieved	Planned	Achieved
Baluchistan	251	251	251	253
KPK	323	239	323	323
Punjab	538	538	538	537
Sindh	616	616	616	616
AJK	300	272	272	271
Total	2028	1916	2000	2000
Achievement (percent)	NA	95 percent	NA	100 percent

Data Collection and Fieldwork

The UCs selected in the survey were mostly located in urban areas. However, there were some UCs in the treatment districts which were located in the rural areas also. Accordingly, some UCs in the control group were selected from the rural areas, too. Considering the differences in the geographical, demographic and socio-cultural makeup of urban and rural areas of Pakistan, different methods were adopted for collecting data in urban and rural areas. These are described in detail below.

Field Methodology – Urban

In the baseline for each UC an area list was developed, next a starting point was selected in each area. Starting points were famous landmarks, market, plaza, or important place that could be considered the centre of the area. Five interviews were conducted around each starting point. This provided a maximum geographical spread within the area. Households were selected using a ballot method where

numbers 1–9 were written on different pieces of paper and put in a box. The surveyor picked one piece of paper and started the interview from the house corresponding to the number on the paper. Since the target respondents for this survey were females, the Right Hand Rule was used to cover the area. This allowed a systematic procedure to minimize surveyor's bias. Further, once a household was selected, screening criteria were used and only houses with children between the ages of 0–59 months were selected. In cases where there was more than one mother in a household, the mother with the most recent birthday was selected.

In the endline, interviews were conducted in the same UC using the baseline contact sheets. The first household interviewed was the same as the first successful interview in the baseline. The same skipping strategy was used as in the baseline so that interviews were conducted in every third household. In situations where more interviews need to be conducted because of refusal by respondents, the standard random walk procedure was used.

Let's suppose there was a rejection in the last household which was covered in the baseline and 2 more interviews were remaining. We contacted the next household and if we had a successful interview there we skipped two households and contacted the household for the remaining interview. In case there was a rejection from the next household also, we went to the next household until we had a successful interview. During endline also, a total of 10 interviews were conducted around a starting point. For the selection of a household around a particular starting point Right Hand Rule[18] was used.

Field Methodology – Rural

Each village was divided into four hypothetical quarters. a starting point was selected in each quarter, and a total of three interviews were conducted around each starting point. Two households were skipped after one successful interview. As with the case in the urban survey, only households with children between the ages of 0-59 months were selected.

In the endline survey the same villages which were selected in baseline were selected, and the same field method was used. Each village was divided into four hypothetical quarters that were approximately the same as those in the baseline. Next a baseline household was used as the starting point in each quarter and four interviews were conducted in each quarter.

For the selection of a household around a particular starting point Right Hand Rule was used.

[18] *Right Hand Rule* - According to the right hand rule, the households were selected on the right side, and by taking right turns until the desired number of interviews were completed. Basic rationale behind adopting this method is to prevent any kind of bias from surveyors' side.