INTERNATIONAL
FOOD
POLICY
RESEARCH
INSTITUTE

A REPORT ON THE SAMPLE SIZES USED FOR THE EVALUATION OF THE EDUCATION, HEALTH, AND NUTRITION PROGRAM (PROGRESA) OF MEXICO

by

Jere R. Behrman and Petra E. Todd

January 25, 1999 (revised version of January 19, 1999)

INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE
Food Consumption and Nutrition Division
2033 K Street N.W.
Washington, D.C. 20006, U.S.A

ACKNOWLEDGMENT

This Research Report is part of the PROGRESA Evaluation project of IFPRI. Behrman and Todd worked on this report as consultants to the PROGRESA Evaluation project of IFPRI. Behrman and Todd alone, and not IFPRI or PROGRESA, are responsible for all contents of this report. Correspondence regarding this paper should be sent to Jere R. Behrman, Economics, McNeil 160, 3718 Locust Walk, University of Pennsylvania, Philadelphia, PA 19104-6297, USA; telephone 215 898 7704, fax 215 898 2124, e-mail jbehrman@econ.sas.upenn.edu or to Petra Todd, Economics, McNeil 160, 3718 Locust Walk, University of Pennsylvania, Philadelphia, PA 19104-6297, USA; telephone 215 898 4084, fax 215 573 2057, e-mail ptodd@econ.sas.upenn.edu.

1. INTRODUCTION

The aim of PROGRESA is to provide support for families living in conditions of extreme poverty in small communities in rural Mexico in order to broaden their opportunities and capabilities to attain higher levels of well-being. The Program carries out actions conducive to raising their living standards by improving opportunities for education, health and food.

Specifically, PROGRESA has the following objectives:

- To substantially improve the conditions of education, health and nutrition of poor families, particularly children and their mothers, by providing sufficient quality services in the areas of education and health, as well as providing monetary assistance and nutrition supplements.

- Integrate these actions so that educational achievement is not affected by poor health or malnutrition in children and young people, or because they carry out work that makes school attendance difficult.

- Ensure that households have sufficient means and resources available so that their children can complete their basic education.

- Encourage the responsibility and active participation of parents and all family members in improving the education, health and nutrition of children and young people.

- Promote community participation and support for the actions of PROGRESA, so that educational and health services benefit all families in the localities where it operates, as well as uniting and promoting community efforts and initiatives in actions that are similar or complementary to the Program.

PROGRESA is made up of three components that are closely linked to each other:

- Educational grants to facilitate and encourage the educational aspirations of children and young people by fostering their enrollment and regular school attendance, and promoting parents' appreciation of the advantages of their children's education. At the same time, actions will be carried out to improve the quality of education.

- Basic health care for all members of the family and strengthening the quality of services as well as reorienting individuals and health services towards taking preventive actions towards health care and nutrition.

- Monetary transfers and nutrition supplements to improve the food consumption and nutritional state of poor families, emphasizing that the purpose of this is to improve the family's food intake, particularly of children and women, who are generally the members of households who are perceived to suffer most from nutritional deficiencies.

Good evaluation of a program such as PROGRESA requires a systematical approach that integrates closely underlying models of the behaviors of those to whom the project is directed, data on those behaviors, and estimation of program effects based on those models and on those data. PROGRESA recognized the importance of collecting good data as part of the evaluation and therefore undertook a baseline survey in March 1998 (ENCEL98M) that is to be followed up with subsequent surveys at approximately half-year intervals. This survey was conducted in 320 randomly selected treatment localities (in which treatment was initiated soon after the baseline survey) and 186 control localities (in which there has been no treatment to date). All 506 of these localities were selected in a stratified random selection procedure (with stratification by populations of localities) from the localities identified by PROGRESA as being eligible to participate in the program because of a "high degree of marginality" that was determined primarily on the bases of analysis of data in the 1990 and 1995 population censuses (1990 Censo, 1995 Conteo). Within localities considered eligible for treatment individual households also were identified as being eligible for treatment ("poor"[1]) based in part on discriminant analysis of a 1997 census that PROGRESA conducted in the localities that had been determined to be eligible for participation in the program (ENCASEH).[2]

There are a number of different aspects of the identification of beneficiary localities and of beneficiary households and of the treatment and control groups that are being evaluated as part of the IFPRI overall evaluation of PROGRESA.[3] This research report focuses solely on the question of the sample sizes and how they relate to alternative ways of phasing in the current evaluation control households/localities into treatment. The evaluation sample in ENCEL98M is composed of the 24,077 households that are divided as indicated in Table 1. For certain questions of interest in the evaluation of PROGRESA, however, the household is not the relevant unit. For example for evaluation of some of the educational aspects of the program, what is relevant may be teenage girls in "poor" households.

The question of sample size, thus, has first of all to do with what size of program effects may be identified with the PROGRESA evaluation sample and how these may differ depending on the aspect of the program under consideration and the nature of the targeted sample. But beyond this

[1] Because PROGRESA is aimed at small poor localities a high proportion of the households in these localities may be poor by many criteria. We use "poor" throughout this report in quotation marks to refer to those household classified as "poor" by PROGRESA.

[2] More details about the PROGRESA selection procedures are provided in PROGRESA (n.d. a, b) and in Behrman, Davis, Levy and Skoufias (1998).

[3] Behrman, Davis, Levy and Skoufias (1998), for example, consider the targeting of beneficiary households. Other forthcoming project research reports will consider the targeting of beneficiary localities and the randomness of the assignments to the control and treatment groups for the evaluation sample.

set of questions, there are further important questions about phasing in the initial control group for treatment. First of all there is the basic question of what are the implications for evaluation of a reduction in the size of the remaining control group. Secondly, as discussed at the IFPRI-PROGRESA workshop on 18 November 1998, there is the question of whether it would be possible to enrich the evaluation of individual components of the PROGRESA program by phasing in different subsets of the initial control group with different initial subcomponents of the overall PROGRESA treatment. For example, might it be possible to phase-in the 186 control localities (9,221 control households) into 3-4 groups so that, for example:

- Group A is the ongoing control group.
- Group B receives educational benefits.
- Group C receives the monetary allowance tied to health clinic visits and nutritional supplements.
- Group D receives the full PROGRESA package given currently to all the treatment localities but the benefits are given to males (not mothers).

The partial treatments in B and C are components of the current full PROGRESA program, other components of which could be phased-in to Groups B and D subsequent to the last evaluation survey round so that eventually the individuals in these groups would receive the identical total PROGRESA package. The way in which the treatment in D differs from the standard PROGRESA package, in contrast, is a change in the program that would be valuable to assess one of the maintained assumptions of PROGRESA that there is an advantage in terms of program objectives of directing program resources to women. If the sample sizes would permit such a phasing in, more could be learned about the initial impact of the major components of PROGRESA through analyzing Groups B and C and about the maintained assumption that resources are better directed towards women through analyzing Group D even though once there is full treatment the components of the partial treatments in B and C are bundled together with other PROGRESA program components (conditional on household demographics). This would enrich importantly what could be learned about the effectiveness of PROGRESA program components. Finally we note that for any of these specific treatments it would be possible to explore the impact of changing the treatment levels by randomly assigning different treatment levels to the recipients (e.g., 50% versus 75% versus 100%) of the original treatment. This would be very informative regarding how response are behaviors to the peso level of different treatments. For nonparametric tests, each different treatment level for a particular treatment would require a different subsample (e.g., B50%, B75%, B100%). If this implied too small subsamples more structure could be imposed with parametric tests.

This Research Report considers these sample size questions. Section 2 summarizes the general considerations regarding sample size, treatment effect sizes, significance level, power, and relative size of treatment and control samples. Section 3 gives information on the sample sizes that currently are available in the evaluation sample for questions concerning a number of central aspects of PROGRESA. Section 4 gives the required sample sizes under alternative assumptions

for a number of selected variables from ENCEL98M that are likely to be central in the evaluation of PROGRESA. Section 5 concludes.

2. GENERAL CONSIDERATIONS REGARDING SAMPLE SIZE

In evaluating the PROGRESA program, the main question of interest is how the educational and nutritional components of the program induce changes in the families receiving treatment. Program impacts are measured by comparing treatment group outcomes to those of a control group that consists of families similar to those in the treatment group who did not receive the treatment. When the treatment and control groups are determined by random assignment, average treatment effects can be assessed simply by comparing means of outcome variables of interest in the two groups.[4] For example, to determine the effect of the program on the proportion of children attending school, we can simply compare the proportions attending in the treatment and control group samples. A test of the null hypothesis that the proportions are equal reveals whether there is a discernible effect of treatment. Because the groups are chosen by random assignment, a comparison of means is not subject to selection bias as it could be if the control group were not chosen by random assignment.

In determining whether treatment has an effect, there are two types of tests that are often conducted: (a) test of equality between two means and (b) test of equality between two proportions. Other types of tests may also be of interest, but in this research report we focus on these two common types of tests and the question of how to determine sample sizes needed to carry them out.

First we need to review some general concepts from the statistical testing literature. With any statistical test, there are two types of errors that can be made, known as Type I and Type II errors. Type I error is rejecting the null when it is true and Type II is accepting the null hypothesis when it is false and some alternative hypothesis is true. The probability of making a Type I error is called the significance level of the test. 1 minus the probability of making a Type II error is called the power of the test. It is equivalent to the probability of correctly rejecting the null hypothesis in favor of the alternative when the alternative hypothesis is true. For reference,

- The *level of significance* (α) gives the probability of rejecting the null hypothesis when it is true (type I error).
- The *power* gives the probability of rejecting the null hypothesis when an alternative hypothesis is true (which equals 1 minus the probability of a type II error ($1-\beta$)).

The power and significance are illustrated in figure 1 for a test of the null that the random variable is distributed with mean zero and variance one against the alternative that it is distributed normally with mean one and variance one. The area to the right of the line under the first curve equals the level of significance and the area to the right of the line under the second curve is the power.

[4] Randomization does not allow one to learn about the full distribution of treatment effects without making some additional assumptions, which are not considered in this report. (See Heckman, Smith and Clements 1997).

A desirable property of a test is that it minimizes the probability of making both type I and type II errors (i.e. low significance level, high power). For a given sample size, there is generally a trade-off between significance level and power. A higher significance level comes at the expense of lower power and vice versa. Both α can be decreased and $1-\beta$ increased by increasing the sample size.

In this report, the goal is to determine the sample sizes required to carry out statistical tests that attain a sufficient level of significance and power with regard to variables that will be of interest in the PROGRESA evaluation. If a desired level of significance is specified, as well as the type of test to be carried out, the alternative hypothesis and the desired power, then one can calculate the required sample size. We illustrate how to do this with a simple example of a one-sided test of the equality between two means.

Let y_2 denote the average outcome measure for the treated group and y_1 the average outcome measure for the control group. Suppose we consider the alternative that treatment increased mean treatment group outcomes above control group outcomes by 5%. The null and alternative hypotheses are:

$$H_0: y_1 = y_2$$
$$H_A: y_2 - y_1 = \Delta$$

where we chose Δ to be equal to $1.05\, y_1$. By a central limit theorem, the difference between two sample means, appropriately standardized follows a normal distribution. Let "^" denote the sample mean estimator and let c_α denote the critical value of the standard normal distribution for a α% significance level. The power of the test is equal to

$$1-b = \Pr\left(\frac{\hat{y}_1 - \hat{y}_2 - 0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} > c_a ; H_A \right)$$
$$= \Pr\left(\frac{\hat{y}_1 - \hat{y}_2 - \Delta}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} > c_a - \frac{\Delta}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \right)$$
$$= 1 - \Phi^{-1}\left(c_a - \frac{\Delta}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \right)$$

Where σ_1 and σ_2 are the standard deviations in the outcomes in the control and treatment group samples and Φ^{-1} is the inverse of the standard normal cumulative distribution function. To compute the required sample sizes, n_1 and n_2, we need to specify the desired level of significance (α), the power ($1-\beta$), the alternative (Δ) and the standard deviations (σ_1 and σ_2). Also, , since different combinations of n_1 and n_2 can solve the equation, we need to know the desired ratio of the n_1/n_2 sample sizes

In the next section we give the required sample sizes for a variety of variables in the ENCEL98M data under several plausible alternatives, which correspond to hypothetical treatment effect levels of 5%, 10%, 25% and 50%. We consider two-sided tests of equality between two means and of equality between two proportions. We chose two-sided instead of one-sided tests, because they give more conservative estimates of the required sample sizes and because we wish the allow in testing for the possibility the impact of treatment may be positive or negative. To calculate the required sample sizes, we used the sampsi command available in the STATA software package.

3. SAMPLE SIZES IN THE EVALUATION SAMPLE

As noted in the introduction, the evaluation sample in ENCEL98M is composed of 24,077 households in 506 localities (Table 1). But the relevant sample for comparisons at the household level between treatment and control groups is conditional on the households being classified as "poor", which totals 12,319, with 7,887 (62.6%) in the treatment group of 320 localities and 4,682 (37.4%) in the control group of 186 localities.

For a number of questions of interest in the evaluation of PROGRESA, moreover, the household is not the relevant unit. Instead the relevant sample groups for a number of possible treatment effects, such as those directed towards school-age or preschool age children, are defined by age and possibly by sex (again, conditional on the household being classified as "poor"). Tables 2 and 3, therefore, give the respective sample sizes for children in the 0-5 and 6-16 age groups by gender in "poor" households. Table 2 further divides the sample into those reported sick and not reported sick in the past four weeks because a number of questions in ENCEL98M are asked only about children who are reported sick. Table 3 similarly further divides the sample into those reported enrolled in school and those reported not enrolled in school because a number of questions in ENCEL98M are conditional on whether a child currently is enrolled in school or not. Both of these tables include tabulations by treatment versus control groups.

The data in these tables emphasize that the relevant samples depend importantly on to what extent a question about treatment is age- and sex-specific – and for the former, what ages are relevant. For instance, if the question is poised in terms of effects on all school-age children of both sexes in "poor" households, the sample size is larger than the number of "poor" households because on the average these households have more then one child in the age range 6-16 years: 14,431 in the treatment group and 8,769 in the control group. All else being equal (e.g., significance level, power, standard deviation, ratio of treatment to control sample sizes) this means that smaller percentage treatment effects can be discerned for all school-age children from "poor" households than for all "poor" households. On the other hand if questions are posed about treatment effects that are in terms of narrowly defined age and sex groups, the sample sizes are very small. For instance there are only 33 boys and only 38 girls who are nine years old and not enrolled in school in the evaluation sample for "poor" households. For such sample sizes only very large treatment effects are discernible *ceteris paribus*. Given the age-specific enrollment patterns in Table 3, for example, this means that questions regarding treatment effects among those who are not enrolled among pre-teens at least are likely to be explorable only by combining across ages and/or sex.

For the age ranges of particular interest to the educational component of PROGRESA – the 12-16 age range – the numbers not enrolled are much larger so some interesting analysis of those not enrolled may be possible by age defined in terms of a year (or some other small range) and/or by sex.

One other group on which the ENCEL98M questionnaire places some emphasis is currently pregnant women (questions 99-107 and the Annex for pregnant women). Based on the answer to question 99 there are 644 respondents (388 in treatment households, 256 in control households) in "poor" households who were claimed that they were pregnant at the time of the survey, plus 27 other women in "poor: households (15 in treatment, 12 in control households) who also claimed that they were pregnant and answered the questions in the questionnaire Annex.

4. SAMPLE SIZES NECESSARY TO DISCERN TREATMENTS EFFECTS OF VARIOUS SIZES FOR DIFFERENT TREATMENTS

In this section we indicate what sample sizes would be needed to discern treatment effects of given percentages for a number of selected variables from ENCEL98M. We select these variables to include what we understand are some of the major outcomes of interest from the point of view of the objectives of PROGRESA (as summarized in Section 1 above) though in some cases with an illustrative example for a class of questions (e.g., the proportion of chicken consumed that is home-produced as an example of similar questions about a number of food products in question 35). Table 4 presents sample sizes necessary to discern the indicated treatment effects in percentage terms for 33 variables.

Each of these sample sizes for individual variables is conditional on a number of assumptions that are discussed in Section 2: the significance level (the default of 0.05 is used), power level (the default of 0.10 is used), whether the test is one-sided or two-sided (as noted in Section 2 we adopt the conservative assumption that two-sided tests are appropriate because, given the multiple dimensions of the PROGRESA treatment, treatment effects might be in either direction for a particular outcome), the mean and the standard deviation of the outcome variable (the ENCEL98M values for "poor" households are used; the means are given in Table 4 for all the variables and the standard deviations are given for the variables that are not proportions[5]), the percentage treatment effect to be discerned (5%, 10%, 25% and 50% increases are used[6]) and the

[5] For any variable that is a proportion the standard deviation of course is a simple function of the proportion (i.e., square root of the product of the proportion and one minus the proportion).

[6] Two comments are appropriate: First, for some of the proportional variables increases of 25% and/or 50% are not possible because 1.0 is the upper limit. In these cases we do not have entries in the table. Second, for most of the variables our *prior* is that the treatment effects are likely to be positive, so we consider only positive changes for all of the variables (even though there are three proportions – P007 not in school because of resource constraints, P011 expected grade completion level no more than primary, P056 was child sick in last 4 weeks -- plus one level variable, P057 duration of sickness for which our *prior* is that the direct treatment effect is likely to be negative). Which direction of change is hypothesized changes the required sample sizes

ratio of the control to the treatment sample size. For the ratio of the control to the treatment sample size four alternatives are used:

> 1.0 (what would result if the current control group were subdivided into equal groups and comparisons made, for example, between the ongoing control group A and any one of the others – B, C, or D – in the example given in Section 1),
>
> 0.6204 (the ratio of the initial number of control households to the initial number of treatment households in ENCEL98M),
>
> 0.3102 (the ratio of half of the initial number of control households to the initial number of treatment households in ENCEL98M -- which would be relevant if half of the initial control households were phased into PROGRESA treatment and comparisons were made between the remaining control households and the initial treatment group),
>
> 0.1551 (the ratio of one quarter of the initial number of control households to the initial number of treatment households in ENCEL98M -- which would be relevant if three-quarters of the initial control households were phased into PROGRESA treatment and comparisons were made between the remaining control households and the initial treatment group).

Table 4 gives, for each variable considered, 16 groups of estimates of required sample sizes – for each of the ratios of the control to the treatment sample size considered, there are four groups of estimates for each of the four percentage treatment effects (5%, 10%, 25% and 50%) considered. For each variable there are two rows, with the first row giving the necessary size of the treatment sample and the next row giving the necessary size of the control sample (which are the same, of course, for the first four columns in which the ratio of the control to the treatment sample is one).

We group the variables in Table 4 into those from the Education Module, Individual Level Variables (Part 1), Education and Consumption Module, Household Level Variables (Part 2), Children's Health Module, Individual Variables (Part 3), Household and Women's Health Module (Part 4), and Women's Status Module (Part 5). These are in the same order as are the questions in ENCEL98M. We give in Table 4 the particular question number in ENCEL98M for each variable (or the questions from which the variable was constructed for constructed variables). Considerable detail is provided in this table for interested readers. We here note five general aspects the patterns in the required sample sizes that are illustrated in the table.

> 1. To reduce the size of a discernible treatment effects by equal percentage increments

---

because it changes the variances, particularly for proportions near the limits of zero and one, (though for the proportions that we consider, not by an order of magnitude). For example, for the proportion that we consider that is closest to the limits – P137 women and men have the same rights ($p = 0.883$) and for a ratio of control to treatment of 0.6204, to detect a 10% change the required sample size is 536 for an increase and 796 for a decrease.

requires proportionately larger increases in the sum of the two sample sizes (i.e., treatment plus control) as the discernible percentages become smaller. For example, for the first variable in the table (P006 proportion attending school) for the ratio of control to treatment sample sizes in the initial evaluation sample (0.6204) it is necessary to add an average of 28 observations for every 5% reduction in discernible treatment effects between 50% and 25% (with the sum increasing from 64 to 206 for a 25% reduction in discernible treatment effects), to add an average of 158 observations for every 5% reduction in discernible treatment effects between 25% and 10% (with the sum increasing from 206 to 681 for a 15% reduction in discernible treatment effects) and to add 237 observations for the 5% reduction in discernible treatment effects between 10% and 5% (with the sum increasing from 681 to 3018 for a 5% reduction in discernible treatment effects). For a number of variables (45% of those considered in the table) this means that even with the initial evaluation samples treatment effects as small as 5% could not be identified.

2. The required sum of the two sample sizes increases as the ratio of control to treatment sample sizes departs further from one (in our case with the control group smaller than the treatment group because of the design of the PROGRESA evaluation sample), with this increase small for small divergences from one but becoming fairly large for larger divergences from one. For example, for the same variable considered in point 1 (P006 proportion attending school) for a treatment impact magnitude of 10%, the required sum increases from 636 (with the ratio equal to one) to 681 (with the ratio equal to 0.6204) to 901 (with the ratio equal to 0.3102) to 1404 (with the ratio equal to 0.1551).

3. The required sample size for the smaller of the two groups (in our case the control group) nevertheless declines a fair amount (though not as rapidly in absolute terms as the sample for the other group increases) as the ratio departs from one. With regard to the same example as in points 1 and 2, the required sample size for the smaller of the two groups decreases from 318 (with the ratio equal to one) to 261 (with the ratio equal to 0.6204) to 214 (with the ratio equal to 0.3102) to 189 (with the ratio equal to 0.1551). This means that for many variables for many comparisons between the initial treatment sample and the modified (with phasing in of treatment) remaining control sample, the latter could be much smaller than the initial control sample size given the relatively large size of the initial control sample.

4. For the variables that are defined as proportions, the required total sample size decreases as the proportion p increases. Table 5 illustrates this point by giving the required total sample sizes for discerning a 10% treatment impact with a control/treatment ratio of 0.6204 (as in the initial evaluation sample) for variables that are ordered in terms of decreasing proportions. These range from 22,287 for P011 expected grade completion level no more than primary (p=0.1730) to 436 for P137 women and men have the same rights (p=0.8830), with intermediate values such as 4,436 for P007 not in school because of resource constraints (p=0.5044). These differences are considerable. They reflect the combination of two effects. First, the variance (=p(1-p)) is larger the closer to 0.50 is p, so that for this reason alone increased sample sizes are required as p increases from 0 to

0.50 (but above 0.50 the variance declines as p is increased further). Second, a given percentage change implies a larger proportional change as p increases, which implies a smaller required sample size (and which outweighs the former effect for $p > 0.5$).

5. For the level variables, the required sample sizes increase *ceteris paribus* as the variance increases for a given mean. Consider, to illustrate, the required total sample sizes for discerning a 10% treatment impact with a control/treatment ratio of 0.6204 (as in the initial evaluation sample) for the expenditure variables with means of the same order of magnitude (in the range of 35 to 51 pesos per time period) but with variances that differ relatively greatly (in the range from 180 to 10032) as summarized in Table 6. The required sample sizes increase monotonicaly with the variances for this set of variables, ranging from 441 to 17289. If both the means and the variances vary, however, the required sample size changes in a more complicated way, and is not simply monotonicaly related to the variance nor to the coefficient of variation.

What do these calculations of required sample sizes mean for what percentage treatment impacts can be estimated with the PROGRESA initial evaluation data and alternative procedures for phasing in part of the initial control group to treatment status? Table 7 provides a basis for answering this question. This table is parallel to Table 4 in organization. But the entries combine the information in Section 3 about actual sample sizes in the initial PROGRESA evaluation sample for different subpopulations of interest with the minimum required sample sizes in Table 4. Each entry indicates whether, under the assumptions for that cell in the table, a treatment effect of that percentage magnitude can (Y) or cannot (N) be discerned. For the age-specific data for children in Parts 1 and 3 of the table, moreover, the cell indicates how many years have to be combined if the sample size is not large enough for single ages (and likewise whether both sexes have to be combined for the 6-16 year age range). Because the ages of particular interest, given the PROGRESA objectives, are the older ages within the 6-16 year range for schooling, the entries in the table indicate how many years (if more than one) starting backwards with 16 are required in Part 1 of the table on schooling.[7] Likewise, because the health of infants and very young children is of particular interest, the entries in the table indicate how many years (if more than one) starting with zero (less than one) are required in Part 3 of the table. We now summarize the implications of this table for what percentage impacts can be discerned for the initial evaluation sample and for two alternatives for phasing-in for treatment part of the initial control sample.

Implications for initial evaluation sample: The initial evaluation sample has a control-to-treatment ratio of 0.6204 so the second group of four columns in Table 7 (and in Table 4) is relevant. The table indicates that for 28 of the 33 included variables (85%) the actual samples are large enough to discern a treatment impact of 10% and for 18 of the 33 variables (55%) the actual

[7] That is, the first entry in the first row of Part 1 of the table ("Y (9yrs,both)") means that the required sample size can be obtained only if the data for ages 8-16 for both sexes are combined and the sixth entry in this row ("Y (both)") means that the required sample size can be obtained for 16 year olds only if both sexes are combined.

samples are large enough to discern a treatment impact of 5%. For four of the five variables for which a 10% impact could not be discerned – P007 not in school because of resource constraints, P011 expected grade completion level no more than primary, P05003 how much did you spend on clothes for girls, and P129 if the woman has extra income she should decide what to do with it – the sample sizes are large enough to discern a 25% treatment impact. For the fifth variable for which a 10% impact could not be discerned – P048 how much did you spend on alcohol- the sample size is sufficiently large to discern a 50% treatment impact. For most of the schooling effects the sample sizes are sufficiently large to discern a 10% treatment impact only for samples combined across ages and sex. Thus the initial evaluation sample is large enough to detect a 10% treatment effect on most of these selected variables (subject to the caveat about combining ages and sexes for the schooling effects) – and for most of the other outcome variables that might be considered in the project. But a 5% impact could be discerned only in little more than half of these variables and there are some variables for which the proportions are so small or the variances so large that even a 25% treatment impact could not be discerned with the initial evaluation sample. We also note here that for use of the initial evaluation sample, as for the alternative means of phasing in part of the initial control sample that are discussed below, a maintained assumption is that members of the control sample who know (or expect) that they eventually will receive treatment do not change their behaviors in anticipation of that treatment prior to the initiation of treatment.

Implications for phasing-in for treatment half of the initial control sample: If the initial control sample were cut in half, what implication does this have for the required sample sizes for comparing the original treatment sample with the remaining half of the original control sample? The third group of four columns in Table 7 (and Table 4) refer to this comparison. The table indicates that for 21 of the 33 included variables (64%) the samples would be large enough to discern a treatment impact of 10% and for 14 of the 33 variables (42%) the samples would be large enough to discern a treatment impact of 5%. For all six of the variables for which a 10% impact would be discernible in the original evaluation sample but not if half of the initial control were phased-in to treatment and not used for evaluation, a 25% treatment effect would be discernible using only half of the initial control sample for the ongoing control group.

The summary in Table 7 does not use the half of the control sample that is phased-in. Some illustrative possibilities for using these observations include: (1) They could be added to the treatment sample if they are given the full treatment. But if the treatment effect changes over time measured from the inception of treatment, it is important that these observations be combined with the original treatment data with sensitivity to the lapsed time since treatments were initiated (which should be possible in principle but which might be confounded by seasonality). If these data were added to the treatment sample they would increase that sample by almost 30% (0.5*4682/7887 from Table 1) which would increase the possibility of discerning small effects. (2) They might be given the same partial treatment such as for one of the alternatives B-D discussed in Section 1. If this entire group were given the same partial treatment, comparisons could be made with (a) the original treatment group and (b) with the remaining control group. (3) They might be split, say, into two equal partial treatment groups. In such a case comparisons could be made such as those that are discussed for the next alternative.

Implications for phasing-in for treatment three-quarters of the initial control sample: If the initial control sample were cut in a fourth, what implication does this have for the required sample sizes for comparing the original treatment sample with the remaining quarter of the original control sample? The fourth group of four columns in Table 7 (and Table 4) refers to this comparison. The table indicates that for 19 of the 33 included variables (58%) the samples would be large enough to discern a treatment impact of 10% and for 10 of the 33 variables (30%) the samples would be large enough to discern a treatment impact of 5%. For all nine of the variables for which a 10% impact would be discernible in the original evaluation sample but not if three-quarters of the initial control were phased-in to treatment and not used for evaluation, a 25% treatment effect would be discernible using only one quarter of the initial control sample for the ongoing control group.

The summary in the last group of four columns in Table 7, again, does not use the three-quarters of the control sample that is phased-in. One possibility, again, is that these observations could be added to the treatment sample if they are given the full treatment (with the same qualifications as given above about sensitivity to the lapsed time since treatment was initiated and seasonality). If these data were added to the treatment sample they would increase that sample by almost 45% (0.75*4682/7887 from Table 1) which again would increase the possibility of discerning small effects beyond that indicated by the last four columns in Table 7.

These phased-in former controls, again, might instead be given the partial treatment such as for the alternatives B-D discussed in Section 1. If this group were split into three equal partial treatment groups, two comparisons would seem natural. Both of these comparisons would be informative about the impact of the initial stages of full versus specific treatment under the maintained assumption that members of the control sample who know (or expect) that they eventually will receive full treatment do not change their behaviors in anticipation of that full treatment prior to the initiation of full treatment (and in particular, when some specific treatment is initiated in isolation). First, each of these groups would be equal to one quarter of the original control sample, so the last four columns in Table 7 are applicable for comparing the partial phase-in of the PROGRESA treatment with the full treatment experienced by the original treatment sample (with the same qualifications once again as given above about sensitivity to the lapsed time since treatment was initiated and seasonality). Second, comparisons could be made between the quarters of the original control sample that receive partial treatment with the quarter of the original control sample that continues to be a control group to investigate the initial impact of some specific partial treatment in comparison with (for the time being) no treatment. For these comparisons the caveat above about sensitivity to the lapsed time since treatment was initiated and seasonality is not relevant. For such comparisons the ratio of the treatment to the control group is one, so the first four columns in Tables 4 and 7 (the latter because the appropriate sample size numbers were used) are relevant. The first four columns of Table 7 indicate that for 18 of the 33 included variables (55%) the samples would be large enough to discern a treatment impact of 10% and for 9 of the 33 variables (27%) the samples would be large enough to discern a treatment impact of 5%. For all ten of the variables for which a 10% impact would be discernible in the original evaluation sample but not if one quarter of the initial control with some specific phased-in treatment were compared with one quarter of the initial control sample still used as a control, a 25% treatment effect would be discernible using only one quarter of the initial control

sample for the ongoing control group.

Implications for reducing the original treatment sample given that the initial control sample is cut in half: Above we have discussed options of reducing the ongoing control sample while maintaining the initial treatment sample. We consider some implications of different ways of phasing in part of the control sample to full or (initially) partial treatment.

Another possibility is to reduce the size of the ongoing treatment sample by randomly selecting a subset of that sample to be continued in the treatment sample. This would reduce the data cost for collecting future rounds of data for evaluation, which would increase resources that could be used for special surveys and special modules as part of the evaluation. This possibility raises the question, what are the tradeoffs from reducing the ongoing treatment sample.

Tables 8 and 9 summarize the implications of reducing the ongoing treatment sample under the maintained assumption that the ongoing control sample will be half of the initial control sample. These tables summarize, from left to right, four options:

- reducing the ongoing treatment sample to the same size as half of the initial control sample so that the ratio of the ongoing control sample to the ongoing treatment sample is one (which implies reducing the ongoing treatment sample to 0.310 of the original treatment sample).

- reducing the ongoing treatment sample to half of the initial treatment sample so that the ratio of the ongoing control sample to the ongoing treatment sample is 0.6204.

- reducing the ongoing treatment sample to three quarters of the initial treatment sample so that the ratio of the ongoing control sample to the ongoing treatment sample is 0.4138.

- maintaining the ongoing treatment sample at the size of the full initial treatment sample so that the ratio of the ongoing control sample to the ongoing treatment sample is 0.3102.

Tables 8 and 9 are parallel to Tables 4 and 7. The fourth option in Tables 8 and 9 is identical to the third option in Tables 4 and 7 in which the ongoing treatment sample is the same as the initial treatment sample and the ongoing control sample is half of the initial control sample, but is repeated for ease of comparison.[8]

Table 9 indicates that the samples would be large enough to discern a 10% impact for 64% of the

[8] In fact three of the groups of columns in Table 8 are identical to three of the groups of columns in Table 4 because these two tables give the required treatment and control groups for different ratios of the control to the treatment samples, and several the same ratios are useful for illustrating the required sample sizes whether the ongoing control sample is reduced below the initial control sample or the ongoing treatment sample is reduced below the initial treatment sample. But only one group of columns is the same in Tables 7 and 9.

included variables with the full initial treatment sample and half of the initial control sample or with three-quarters of the initial treatment sample and half of the initial control sample and for 58% of the included variables with half of the initial treatment sample and half of the initial control sample or with the ongoing treatment and control samples each equal to half of the initial control sample. Table 9 also indicates that the samples would be large enough to discern a 5% impact for 42% of the included variables with the full initial treatment sample and half of the initial control sample, for 36% of the included variables with three-quarters of the initial treatment sample and half of the initial control sample and with half of the initial treatment sample and half of the initial control sample, and for 33% of the included variables with the ongoing treatment and control samples each equal to half of the initial control sample. Thus, as noted above, whether or not the ongoing treatment sample is reduced from the initial treatment sample, for many variables identification of a 5% impact is not possible. But there would be a relatively limited cost in terms of identifying a 10% impact of randomly reducing the ongoing treatment group to the sizes considered in Tables 8 and 9.

5. CONCLUSION AND SUMMARY

We have evaluated options for reducing the control and the treatment samples sizes from their initial levels. Before turning to a summary of our results, it is important to emphasize the following four points.

(1) Our evaluation depends on the assumption that the original assignments to the control and treatment groups were random, an assumption that will be explored in a subsequent research report.

(2) Our evaluation depends on any households that are dropped from the initial control or initial treatment samples be dropped randomly.

(3) Our evaluation considers the overall national sample, not regional subsamples. If regional or some other more disaggregated subsamples are of interest, larger national samples are required *ceteris paribus* so that the regional subsamples are sufficiently large.

(4) To evaluate some of the individual schooling and child health effects it will be necessary to combine across ages and/or sexes. The need to do so increases as the sizes of the ongoing control and ongoing treatment samples decrease. In our analysis of Section 4 we indicate that evaluation of a given percentage impact is possible if it is possible by combining across the age/sex groups in the sample. But such combinations may be costly for evaluating some of PROGRESA's objectives. If, for example, the impact on teenage female school attendance is of great interest, impacts that can be discerned only by combining males and females for ages 8-16 are not very useful.

For these reasons we adopt a conservative position in our summary that follows by being cautious about recommending more than limited reductions in ongoing control and treatment sample sizes.

The original evaluation sample sizes for treatment and control groups, under conservative assumptions (e.g., using two-tailed tests) permit testing for 10% impact of treatment on a high proportion (85%) of the variables considered in this report and for a 5% impact on the majority (55%) of these variables if groups are combined across ages for some of the schooling and health variables. Phasing-in some of the control sample for treatment has some cost in terms of being able to evaluate small treatments and in terms of needing to combine across more ages for some of the schooling and child health outcomes, but still a 10% impact could be discerned for 64% of the variables considered if the initial control group were cut in half and for 58% of the variables if the control group were reduced to one quarter of the original size. For all of the variables for which a 10% impact could be discerned with the original evaluation sample but not with the control group reduced to half or to a quarter of its original size, moreover, 25% effects could be discerned with these particular smaller control samples.

Therefore the PROGRESA commitment to initiate phasing-in treatment for part of the initial control sample does not seem too costly in terms of the potentiality for evaluating effects of the magnitude of 10-25% (though the prospects are limited in any case for evaluating smaller effects on many outcomes). It is important, of course, that whatever control groups are phased-in are selected randomly (though possibly with the stratification of population of localities that was used for the original selection of the evaluation sample).

That still leaves open the questions of (1) what share of the original control group should be phased-in before the completion of collection of all the evaluation data rounds and (2) what treatment (partial or full) should be given to those initial control households/localities that are phased-in before the completion of collection of all the data rounds. There is some tension between the answers to these two questions. From a conservative position regarding the potentiality for evaluation of the basic PROGRESA package, with regard to question (1) it would be desirable to keep the ongoing control sample as large as possible, and half of the original control sample would be better than a quarter. On the other hand the phasing-in of part of the control group with partial treatments such as only educational benefits or only monetary allowances tied to health clinic visits and nutritional supplements or with the benefits given to fathers rather than mothers is the only possibility within the whole evaluation project to assess any of the components of the overall package separately. And more such partial treatment alternatives can be explored with useful sample sizes for such partial treatments if more of the initial control group is phased-in with partial treatments.

Based on these competing considerations, our judgment is that it would be desirable (a) to adopt of conservative strategy by maintaining 50% of the original control group as a ongoing control for the evaluation of the overall package and (b) to use the rest of the original control group to phase-in two specific partial treatments – namely B (educational benefits only) and C (health/nutrition benefits only) immediately (before the next evaluation round) for a randomly selected set of localities that are in the initial control group. We do not recommend D (giving the benefits to males rather than females) because we perceive the cost of so doing in terms of sample size would be too great relative to the expected gains regarding what would be learned that could change how the PROGRESA program evolved even though we think that it would be of

considerable general interest to learn whether the sex of the recipient mattered.[9] We also recommend that B and C be introduced initially with different peso transfers (say, 50%, 75% and 100% of the basic PROGRESA program amounts) assigned randomly by localities (among those picked to receive the particular specific treatment). This would permit evaluation of how responsive behaviors are to changing the amount of the benefits received. The cost of doing so is that the tests would have to be parametric because of the small sample sizes for each level of treatment rather than nonparametric. In our judgment the gain in learning about what the behavioral responses are to changing the magnitude of treatment benefits would be worth this cost. We emphasize that it is important that the program implementation for B and C (whether or not different benefit levels are administered) be identical in all respects to the delivery of these program components as part of the overall PROGRESA package in other localities in order to be able to ascertain what is the impact of the specific treatments per se.

We think that following the recommendations in the previous paragraph would enhance considerably what could be learned from the evaluation of PROGRESA regarding specific program components.

A related question is what would be the cost of reducing the ongoing treatment sample. The calculations at the end of Section 4 indicate that the original treatment sample could be reduced by a quarter or by half or possibly more with limited impact of the probabilities of being able to identify impacts of 10%. There basically are but limited gains in having the ongoing treatment sample be much bigger than the ongoing control sample. If there is a substantial gain in other information that could be obtained if resources were made available through reducing the ongoing treatment sample, then this option should be given serious consideration. Possible gains might include the collection of more information in each evaluation round survey or collecting related special information with additional surveys between the evaluation rounds. But it seems desirable to be conservative once again in reducing the ongoing treatment sample size for all the reasons that we discuss at the start of this section plus the fact that reducing the ongoing treatment sample also reduces the probabilities of discerning impacts in the subsamples of the original control group that are phased in if our recommendation above is followed. Therefore we recommend that such a reduction in the ongoing treatment sample be made only if there is perceived to be large gains in terms of data collection for other aspects of the evaluation. If there are perceived to be such potential large gains, we still would take a conservative position that the original treatment sample should be reduced randomly by no more than a third.

[9] We are presuming here that the administrative cost of identifying the program recipients by sex is not all that great. Were it large, it might be desirable even from the point of view of possible future evolution of PROGRESA to include D as an alternative to see whether the additional gains in attainment of the overall PROGRESA objectives warranted the additional cost.

References

Behrman, Jere, Benjamin Davis, Dan Levy, and Emmanuel Skoufias, 1998, "A Preliminary Evaluation of the Selection of Beneficiary Households in the Education, Health and Nutrition Program (PROGRESA) of Mexico," Washington, DC: Research Report, PROGRESA Evaluation Project of IFPRI, memo.

Heckman, J. J., J. Smith, and N. Clements, 1997, "Making the Most Out of Social Experiments," in *Review of Economic Studies*, Vol. 64(4), October.

PROGRESA, n.d., a, "Metodo de seleccion de las localidades PROGRESA," Mexico: PROGRESA, mimeo.

PROGRESA, n.d., b, "Nota Tecnica: Diseno de Meustras Basal y Control," Mexico: PROGRESA, mimeo.

StataCorp, 1997, Stata Statistical Software: Release 5, College Station, Texas: Stata Corporation.

Table 1. Evaluation Sample (ENCEL98M) Numbers of Localities and Numbers of Households Classified by Treatment/Control and "Poor"/"NonPoor"[a]

	Treatment	Control	Total
Localities	320	186	506
Households	14856	9221	24077
	61.7%	38.3%	100%
	100%	100%	100%
Not classified as "Poor"	7019	4539	11558
	60.7%	39.3%	100%
	47.3%	49.2%	48.0%
Classified as "Poor"	7887	4682	12319
	62.6%	37.4%	100%
	52.8%	50.8%	52.0%

[a] Tabulations from ENCEL98M/BaseM. The first percentage under each entry for the household gives the percentage of the row total and the second gives the percentage of the column total.

Table 2
Actual Sample Sizes
Children age 0-5 in Eligible Households (pobre=1)
by Whether Recently Ill, Treatment Status

Age	ill in last 4 weeks[(a)]		not ill in last 4 weeks		Total[(b)]	
	treatment	control	Treatment	control	treatment	control
0	383	207	702	417	1102	633
1	465	272	671	392	1150	685
2	414	228	936	577	1378	821
3	395	249	1025	589	1441	851
4	373	210	1168	698	1568	935
5	334	223	1143	673	1510	916
Total	2364	1389	5645	3346	8149	4841

(a) Based on P056 variable. Excludes nonrespondents.
(b) Includes nonrespondents

Table 3, Part 1
Actual Sample Sizes
Combined Treatment and Control Children in Eligible Households (pobre=1)
by Age (Boys and Girls combined)

Age	enrolled in school[(a)]	not enrolled in school	Total
6	2251	193	2444
7	2281	85	2366
8	2423	77	2500
9	2139	71	2210
10	2313	95	2408
11	2010	90	2100
12	2105	261	2366
13	1461	464	1925
14	1220	677	1897
15	875	844	1719
16	512	753	1265
Total	19590	3610	23200

(a) Based on P006 variable.

Table 3, Part 2
Actual Sample Sizes
Combined Treatment and Control Male Children in Eligible Households (pobre=1)
by Age

Age	enrolled in school[(a)]	Not enrolled in school	Total
6	1157	96	1253
7	1175	37	1212
8	1259	42	1301
9	1076	33	1109
10	1205	48	1253
11	1009	46	1055
12	1148	102	1250
13	795	197	992
14	693	327	1020
15	506	400	906
16	297	403	700
Total	10320	1731	12051

(a) Based on P006 variable.

Table 3, Part 3
Actual Sample Sizes
Combined Treatment and Control Female Children in Eligible Households (pobre=1)
by Age

Age	enrolled in school[(a)]	Not enrolled in school	Total
6	1094	97	1191
7	1106	48	1154
8	1164	35	1199
9	1063	38	1101
10	1108	47	1155
11	1001	44	1045
12	957	159	1116
13	666	267	933
14	527	350	877
15	369	444	813
16	215	350	565
Total			

(a) Based on P006 variable.

Table 3, Part 4
Actual Sample Sizes
Children age 6-16 in Eligible Households (pobre=1)
by Age, Treatment Status (boys and girls combined)

Age	Enrolled in school[(a)]		not enrolled in school		Total	
	Treatment	control	treatment	control	treatment	control
6	1401	847	101	92	1505	939
7	1436	845	54	31	1490	876
8	1547	876	39	38	1586	914
9	1338	801	45	26	1383	827
10	1454	859	48	47	1502	906
11	1228	782	54	36	1282	818
12	1315	790	170	91	1485	881
13	887	574	277	187	1164	761
14	740	480	412	265	1152	745
15	572	303	538	306	1110	609
16	330	182	442	311	772	493
Total	12251	7339	2180	1430	14431	8769

(a) Based on P006 variable.

Table 3, Part .5
Actual Sample Sizes
Male Children age 6-16 in Eligible Households (pobre=1)
by Age, Treatment Status

Age	enrolled in school[(a)]		not enrolled in school		Total	
	treatment	Control	treatment	control	treatment	control
6	732	425	51	45	783	470
7	751	424	23	14	774	438
8	809	450	17	25	826	475
9	688	388	21	12	709	400
10	753	452	28	20	781	472
11	622	387	25	21	647	408
12	723	425	65	37	788	462
13	485	310	113	84	598	394
14	442	251	196	131	638	382
15	340	166	268	132	608	298
16	185	112	237	166	422	278
Total	6530	3790	1044	687	7574	4477

(a) Based on P006 variable.

Table 3, Part 6
Actual Sample Sizes
Female Children age 6-16 in Eligible Households (pobre=1)
by Age, Treatment Status

Age	enrolled in school[(a)]		not enrolled in school		Total	
	treatment	Control	treatment	control	treatment	control
6	672	422	50	47	722	469
7	685	421	31	17	716	438
8	738	426	22	13	760	439
9	650	413	24	14	674	427
10	701	407	20	27	721	434
11	606	395	29	15	635	410
12	592	365	105	54	697	419
13	402	264	164	103	566	367
14	298	229	216	134	514	363
15	232	137	270	174	502	311
16	145	70	205	145	350	215
Total	5721	3549	1136	743	6857	4292

(a) Based on P006 variable.

Table 4, Part 1 (Education Module, Individual Level Variables)
Required Sample Size for Two-sided Test of Equality of Proportions or Equality of Means
significance level α=0.05, power β=0.90
For Different Variables of Interest and Varying Treatment/Control Proportions[1]

Variable of Interest	(n2/n1) ratio=1.0 treatment impact magnitude				(n2/n1) ratio= 0.6204 treatment impact magnitude				(n2/n1) ratio = 0.3102 treatment impact magnitude				(n2/n1) ratio = 0.1551 treatment impact magnitude			
	5%	10%	25%	50%	5%	10%	25%	50%	5%	10%	25%	50%	5%	10%	25%	50%
P006: proportion attending school (p=0.8444)	1418 1418	318 318	*	*	1862 1156	420 261	*	*	3030 940	687 214	*	*	5363 832	1215 189	*	*
P007*[2]: not in school because of resource constraints (p=0.5044)	8328 8328	2096 2096	338 338	83 83	10876 6748	2737 1699	442 275	109 68	17588 5456	4426 1373	715 222	175 55	31011 4810	7803 1211	1259 196	307 48
P008* and P012*[3]: imputed years of education (mean=4.244, sd=2.853)	3799 3799	950 950	152 152	38 38	4961 3078	1241 770	199 124	50 32	8023 2489	2006 623	321 100	81 26	14146 2195	3537 549	566 88	142 23
P009: how is child doing at school? (good=1) (p = 0.6114)	5325 5325	1323 1323	204 204	45 45	6959 4318	1731 1074	268 167	59 37	11262 3494	2803 870	435 135	96 30	19868 3082	4946 768	766 119	166 26
P011: expected grade completion level																
no more than primary (p=0.173)	41202 41202	10552 10552	1805 1805	497 497	53756 33351	13754 8533	2348 1457	645 401	86286 26934	22189 6884	3774 1171	1030 320	152963 23725	39056 6058	6624 1028	1800 280
secondary but not higher (p=0.509)	8103 8103	2056 2056	332 332	81 81	10674 6623	2685 1666	433 269	106 66	17262 5355	4342 1347	700 218	171 54	30496 4721	7655 1188	1234 192	300 47
more than secondary (p=0.318)	18387 18387	4683 4683	787 787	210 210	23997 14888	6108 3790	1025 636	273 170	38775 12029	9862 3060	1651 513	438 136	68328 10598	17367 2694	2902 451	766 119
P014: number of days absent in last 4 weeks (mean=0.4691, sd=1.9588)	397 397	330 330	206 206	113 113	518 322	430 267	269 167	148 92	837 260	695 216	434 135	239 75	1475 229	1226 191	765 119	420 66

1. The treatment impacts considered are always in the direction of increasing the mean or proportion. A * appears when the treatment impact would cause the proportion to exceed 1.
2. Variable constructed from P007. Equals 1 if P007 equals 01, 02, or 03.
3. Variable constructed from P004, P006, P008, and P012. For individuals currently in school, construct Age-6-P012. For individuals not currently in school, construct Age-6-P008 if P008<8, else 0. Recode negative values for this group to 0.

Table 4, Part 2 (Education and Consumption Module, Household Level Variables)
Required Sample Size for Two-sided Test of Equality of Proportions or Equality of Means
significance level α=0.05, power β=0.90
For Different Variables of Interest and Varying Treatment/Control Proportions[(1)]

	(n2/n1) ratio=1.0				(n2/n1) ratio= 0.6204				(n2/n1) ratio = 0.3102				(n2/n1) ratio = 0.1551			
	treatment impact magnitude				treatment impact magnitude				treatment impact magnitude				treatment impact magnitude			
Variable of Interest	5%	10%	25%	50%	5%	10%	25%	50%	5%	10%	25%	50%	5%	10%	25%	50%
P025: Are you or your husband in the school parents association? (p=0.3389)	16708 16708	4251 4251	712 712	189 189	21808 13530	5546 3441	928 576	245 152	35239 10932	8955 2778	1495 464	394 123	62101 9632	15771 2447	2628 408	690 108
P038: How much do you spend on fruits and vegetables? (mean=35.29 sd=31.19)	6567 6567	1642 1642	263 263	66 66	8575 5320	2144 1331	343 213	86 54	13867 4302	3467 1076	555 173	139 44	24451 3793	6113 949	979 152	245 38
P035b01: Home-produced chicken? (for example) (p=0.6641)	4201 4201	1034 1034	154 154	31 31	5493 3408	1354 841	203 126	41 26	8896 2760	2195 681	329 103	64 20	15699 2435	3877 602	581 91	107 17
P038: How much do you spend on food of animal origin? (mean=49.31 sd=40.78)	5750 5750	4069 4069	230 230	58 58	7509 4659	5313 3297	301 187	76 48	12142 3767	8592 2666	486 151	122 38	21409 3321	14149 2350	857 133	215 34
P046: How much money do you use in preparing food in a week? (mean=128.08 sd=73.00)	2731 2731	683 683	110 110	28 28	3567 2213	892 554	143 89	36 23	5767 1789	1442 448	231 72	58 18	10169 1578	2543 395	407 64	102 16
P048: How much did you spend last week on alcohol? (mean=2.054, sd=10.845)	234339 234339	58585 58585	9374 9374	2344 2344	306031 189862	76508 47466	12242 7595	3061 1900	494892 153516	123723 38379	19796 6141	4949 1536	872615 135343	218154 33836	34905 5414	8727 1354
P048: How much did you spend last week on soft drinks? (mean=10.14, sd=15.02)	18444 18444	4611 4611	738 738	185 185	24087 14944	6022 3737	964 599	241 150	38951 12083	9738 3021	1559 484	390 121	68680 10653	17170 2664	2748 427	687 107

(1) The treatment impacts considered are always in the direction of increasing the mean or proportion. A * appears when the treatment impact would cause the proportion to exceed 1.

Table 4, Part 2 continued (Education and Consumption Module, Household Level Variables)
Required Sample Size for Two-sided Test of Equality of Proportions or Equality of Means
significance level α=0.05, power β=0.90
For Different Variables of Interest and Varying Treatment/Control Proportions[1]

Variable of Interest	(n2/n1) ratio=1.0 treatment impact magnitude				(n2/n1) ratio= 0.6204 treatment impact magnitude				(n2/n1) ratio = 0.3102 treatment impact magnitude				(n2/n1) ratio = 0.1551 treatment impact magnitude			
	5%	10%	25%	50%	5%	10%	25%	50%	5%	10%	25%	50%	5%	10%	25%	50%
P05003: How much did you spend in last 6 months on clothes for girls mean=50.80 sd=100.16	32678 32678	8170 8170	1308 1308	327 327	42675 26476	10669 6620	1707 1060	427 265	69011 21408	17253 5352	2761 857	691 215	121682 18873	30421 4719	4868 756	1217 189
P048: How much did you spend last week on medicine? (mean=42.83, sd=13.46)	831 831	208 208	34 34	9 9	1085 674	272 169	44 28	11 7	1754 545	439 137	71 23	18 6	3092 480	773 120	124 20	31 5

I. The treatment impacts considered are always in the direction of increasing the mean or proportion. A * appears when the treatment impact would cause the proportion to exceed 1.

Table 4, Part 3 (Children's Health Module, Individual Level Variables)
Required Sample Size for Two-sided Test of Equality of Proportions or Equality of Means
significance level α=0.05, power β=0.90
For Different Variables of Interest and Varying Treatment/Control Proportions[1]

	(n2/n1) ratio=1.0				(n2/n1) ratio= 0.6204				(n2/n1) ratio = 0.3102				(n2/n1) ratio = 0.1551			
	treatment impact magnitude				treatment impact magnitude				treatment impact magnitude				treatment impact magnitude			
Variable of Interest	5%	10%	25%	50%	5%	10%	25%	50%	5%	10%	25%	50%	5%	10%	25%	50%
P056: Was child sick in the last 4 weeks? (p=0.2955)	20391 20391	5199 5199	879 879	236 236	26611 16510	6780 4207	1145 711	306 190	42996 13338	10945 3396	1843 572	491 153	75763 11751	19273 2990	3238 503	859 134
P057: Duration of sickness (mean=7.0857 sd=5.1458)	4433 4433	1109 1109	178 178	45 45	5790 3593	1448 899	232 144	58 36	9362 2905	2341 727	375 117	94 30	16507 2561	4127 641	661 103	166 26
P060*[2]: Saw a health helper, doctor, nurse or dentist regarding illness (p=0.6593)	4287 4287	1059 1059	159 159	32 32	5605 3478	1387 861	208 130	42 27	9076 2816	2249 698	338 105	67 21	16017 2485	3971 616	597 93	113 18
P068*[3] Is child less than one year old currently being breast fed? (p=0.7492)	2715 2715	654 654	88 88	*	3554 2205	858 533	117 73	*	5764 1788	1395 433	189 59	*	10182 1580	2467 383	332 52	*
P069: Did you take child to health center to determine weight and height? (p=0.8027)	1954 1954	455 455	*	*	2562 1590	599 372	*	*	4161 1291	976 303	*	*	7357 1142	1727 268	*	*
P073-P076: Does child have TB, triple vaccine, polio and measles vaccines? (p=0.8641)	1184 1184	258 258	*	*	1558 967	342 213	*	*	2538 788	559 174	*	*	4494 698	988 154	*	*

I. The treatment impacts considered are always in the direction of increasing the mean or proportion. A * appears when the treatment impact would cause the proportion to exceed 1.
II. Variable constructed from P060.
III. This variable constructed from P068, restricting population to children less than one year old.

Table 8, Part 4 (Household and Women's Health Module)
Required Sample Size for Two-sided Test of Equality of Proportions or Equality of Means
significance level α=0.05, power β=0.90
For Different Variables of Interest and Varying Treatment/Control Proportions

	(n2/n1) ratio=1.0 (half initial control group and equal size treatment group) treatment impact magnitude				(n2/n1) ratio= 0.6204 (half initial treatment group, half initial control group) treatment impact magnitude				(n2/n1) ratio = 0.4138 (3/4 initial treatment group, half initial control group) treatment impact magnitude				(n2/n1) ratio = 0.0.3102 (initial treatment group, half initial control group) treatment impact magnitude			
Variable of Interest	5%	10%	25%	50%	5%	10%	25%	50%	5%	10%	25%	50%	5%	10%	25%	50%
P077: Do you know how to prepare oral serum? (p=0.8223)	1695 1695	389 389	108 108	32 32	2223 1380	513 319	140 87	42 27	2918 1208	675 280	180 75	53 22	3614 1122	838 260	220 69	65 21
P078: Do you boil or put bleach in the water? (p=0.7864)	2176 2176	513 513	64 64	36 36	2852 1770	675 419	84 53	47 30	3740 1548	887 368	111 46	60 25	4630 1437	1099 341	136 43	73 23
P082*[(2)]: Did you or your family go to a medical center to see a doctor in the last year? (p=0.8084)	1876 1876	436 436	115 115	34 34	2460 1527	574 357	148 92	44 28	3228 1336	755 313	191 80	56 24	3997 1240	936 291	234 73	68 22
P102: Did you see a doctor or nurse during this pregnancy? (p=0.5776)	6145 6145	1534 1534	241 241	56 56	8028 4981	2005 1244	316 197	73 46	10505 4347	2625 1087	414 172	96 40	12989 4030	3246 1007	512 159	118 37
P120-P122*[(3)]: Are you currently or have you ever used contraceptives or are you sterilized? (p=0.4168)	11980 11980	3021 3021	499 499	129 129	15640 9704	3942 2446	651 404	168 105	20454 8464	5154 2133	851 353	219 91	25281 7843	6369 1976	1051 327	270 84

(1) Variable constructed from P082 by recoding 1,2, or 3 as went to a doctor.
(2) Constructed variable from P120, P121, P122

Table 4, Part 5 (Women's Status Module)
Required Sample Size for Two-sided Test of Equality of Proportions or Equality of Means
significance level α=0.05, power β=0.90
For Different Variables of Interest and Varying Treatment/Control Proportions[1]

	(n2/n1) ratio=1.0				(n2/n1) ratio= 0.6204				(n2/n1) ratio = 0.3102				(n2/n1) ratio = 0.1551			
	treatment impact magnitude				treatment impact magnitude				treatment impact magnitude				treatment impact magnitude			
Variable of Interest	5%	10%	25%	50%	5%	10%	25%	50%	5%	10%	25%	50%	5%	10%	25%	50%
P129*[1]: If the woman has extra income, she should decide what to do with it. (p=0.1947)	35880 35880	9090 9090	1555 1555	427 427	46814 29044	11850 7352	2022 1255	553 344	75617 23457	19118 5931	3252 1009	885 275	133222 20663	33653 5220	5709 886	1546 240
P137: Women must always obey men (p=0.794)	2068 2068	485 485	59 59	*	2711 1682	639 397	78 49	*	4403 1366	1041 323	125 39	*	7783 1208	1842 286	213 34	*
P137: Women may work out of the household (p=0.708)	3393 3393	826 826	118 118	*	4439 2754	1083 672	156 97	*	7193 2232	1758 546	253 79	*	12700 970	3107 482	446 70	*
P137: Women and men have the same rights (p=0.883)	967 967	202 202	*	*	1274 791	269 167	*	*	2079 645	440 137	*	*	3684 572	774 121	*	*

I. The treatment impacts considered are always in the direction of increasing the mean or proportion. A * appears when the treatment impact would cause the proportion to exceed 1.

Table 5. Required Sample Sizes for Variables Defined as Proportions[a]

Variable	Proportion	Total Required Sample (Control plus Treatment)
P137: Women and men have the same rights.	0.8830	436
P073-P076: Does child have TB, triple vaccine, polio and measles vaccines?	0.8641	555
P006: Proportion children age 6-16 years attending school.	0.8444	681
P077: Do you know how to prepare oral serum?	0.8223	832
P082: Did you or your family go to a medical center to see a doctor in the last year?	0.8084	931
P069: Did you take child to health center to determine weight and height?	0.8027	971
P137: Women must always obey men.	0.7940	1036
P078: Do you boil or put bleach in the water?	0.7864	1094
P068: Is child less than one year old currently being breast fed?	0.7492	1391
P137: Women may work out of the household.	0.7080	1755
P035b01: Home-produced chicken?	0.6641	2195
P060: Saw a health helper, doctor, nurse or dentist regarding illness	0.6593	2248
P009: how is child doing at school? (good=1)	0.6114	2805
P102: Did you see a doctor or nurse during this pregnancy?	0.5776	3249
P011: expected grade completion level secondary but not higher	0.5090	4351
P007: not in school because of resource constraints	0.5044	4436
P120-P122: Are you currently or have you ever used contraceptives or are you sterilized?	0.4168	6388
P025: Parent(s) in school-parents association.	0.3389	8987
P011: Expected grade completion level more than secondary.	0.3180	9898
P056: Was child sick in the last 4 weeks?	0.2955	10987
P129: If the woman has extra income, she should decide what to do.	0.1947	19202
P011: Expected grade completion level no more than primary.	0.1730	22287

[a] From Table 4 with ratio of control to treatment sample of 0.6204 for a 10% treatment impact.

Table 6. Required Sample Sizes for Expenditure Variables with Means of Same Order of Magnitude but Differing Variances[a]

Variable (mean)	Variance	Required Total Sample Size (Control plus Treatment)
P048: How much did you spend on medicine last week? (mean = 42.83)	181.17	441
P038: How much do you spend on fruits and vegetables? (mean = 35.29)	972.81	3475
P038: How much do you spend on food of animal origin? (mean = 49.31)	1663.01	8610
P05003: How much did you spend in last 6 months on clothes for girls? (mean = 50.80)	10032.03	17289

[a] From Table 4 with ratio of control to treatment sample of 0.6204 for a 10% treatment impact.

Table 7, Part 1 (Education Module, Individual Level Variables)
Whether Different Samples Satisfy Required Sample Size for Two-sided Test of Equality of Proportions or Equality of Means
significance level =0.05, power =0.90

Variable of Interest	Quarter of Initial Control Sample and Equal-Size Treatment Phase-in Sample (n2/n1) ratio=1.0 treatment impact magnitude				Initial Evaluation Sample (n2/n1) ratio= 0.6204 treatment impact magnitude				Half of Initial Control Sample and Initial Treatment Sample (n2/n1) ratio = 0.3102 treatment impact magnitude				Quarter of Initial Control Sample and Initial Treatment Sample (n2/n1) ratio = 0.1551 treatment impact magnitude			
	5%	10%	25%	50%	5%	10%	25%	50%	5%	10%	25%	50%	5%	10%	25%	50%
P006: proportion attending school (p=0.8446)	Y (9 yrs, both)	Y (3 yrs, both)	*	*	Y (2 yrs, both)	Y (both)	*	*	Y (4 yrs, both)	Y (both, 2 yrs)	*	*	Y (5 yrs, both)	Y (2 yrs, both)	*	
P007: not in school because of resource constraints (p=0.5044)	N	N	Y (9 yrs, both)	Y (2 yrs, both)	N	N	Y (both)	Y	N	N	Y (2 yrs, both)	Y	N	N	Y (3 yrs, both)	Y (both)
P008 and P012: imputed years of education (mean=4.244, sd=2.853)	N	Y (7 yrs, both)	Y (3 yrs, both)	Y (both)	Y (4 yrs, both)	Y (2 yrs, both)	Y	Y	Y (9 yrs, both)	Y (3 yrs, both)	Y (both)	Y	N	Y (5 yrs, both)	Y (2 yrs, both)	Y
P009: how is child doing at school? (good=1) (p = 0.6114)	N	Y (9 yrs, both)	Y (3 yrs, both)	Y (both)	Y (7 yrs, both)	Y (4 yrs, both)	Y (both)	Y	N	Y (5 yrs, both)	Y (2 yrs, both)	Y	N	Y (6 yrs, both)	Y (2 yrs, both)	Y
P011: expected grade completion level																
no more than primary (p=0.173)	N	N	Y(10 yrs, both)	Y (5 yrs, both)	N	N	Y (4 yrs, both)	Y (2 yrs, both)	N	N	Y (6 yrs, both)	Y (3 yrs, both)	N	N	Y (7 yrs, both)	Y (4 yrs, both)
secondary but not higher (p=0.509)	N	N	Y (4 yrs, both)	Y (2 yrs, both)	Y (10 yrs, both)	Y (4 yrs, both)	Y (2 yrs, both)	Y (both)	N	Y (7 yrs, both)	Y (2 yrs, both)	Y (both)	N	Y (8 yrs, both)	Y (3 yrs, both)	Y (both)
more than secondary (p=0.318)	N	N	Y (7 yrs, both)	Y (3 yrs, both)	N	Y (7 yrs, both)	Y (3 yrs, both)	Y (both)	N	Y (7 yrs, both)	Y (2 yrs, both)	Y (2 yrs, both)	N	N	Y (5 yrs, both)	Y (2 yrs, both)
P014: number of days absent in last 4 weeks (mean=0.4691, sd=1.9588)	Y (5 yrs, both)	Y (4 yrs, both)	Y (3 yrs, both)	Y (2 yrs, both)	Y (2 yrs, both)	Y (2 yrs, both)	Y (both)	Y	Y (2 yrs, both)	Y (2 yrs, both)	Y (2 yrs, both)	Y (both)	Y (3 yrs, both)	Y (3 yrs, both)	Y (2 yrs, both)	Y (2 yrs, both)

N.B. Based on Table 3, using oldest ages in 6-16 age range necessary to get required sample sizes that are indicated in Table 4 (i.e., "2 yrs" means that ages 15 and 16 must be combined, "3 yrs" means that ages 14-16 must be combined). "Both" means that both females and males must be combined or the years doubled. "*" means that the treatment impact would cause the proportion to exceed 1.

Table 7, Part 2 (Education and Consumption Module, Household Level Variables)
Whether Different Samples Satisfy Required Sample Size for Two-sided Test of Equality of Proportions or Equality of Means
significance level =0.05, power =0.90

Variable of Interest	Quarter of Initial Control Sample and Equal Treatment Phase-in Sample (n2/n1) ratio=1.0				Initial Evaluation Sample (n2/n1) ratio= 0.6204				Half of Initial Control Sample and Initial Treatment Sample (n2/n1) ratio = 0.3102				Quarter of Initial Control Sample and Initial Treatment Sample (n2/n1) ratio = 0.1551			
	treatment impact magnitude				treatment impact magnitude				treatment impact magnitude				treatment impact magnitude			
	5%	10%	25%	50%	5%	10%	25%	50%	5%	10%	25%	50%	5%	10%	25%	50%
P025: Are you or your husband in the school parents association? p=0.3389	N	N	Y	Y	N	Y	Y	Y	N	N	Y	Y	N	N	Y	Y
P038: How much do you spend on fruits and vegetables? mean=35.29 sd=31.19	N	Y	Y	Y	N	Y	Y	Y	N	Y	Y	Y	N	Y	Y	Y
P035b01: Home-produced chicken? p=0.6641	N	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	N	Y	Y	Y
P038: How much do you spend on food of animal origin? Mean=49.31 sd=40.78	N	N	Y	Y	Y	Y	Y	Y	N	N	Y	Y	N	N	Y	Y
P046: How much money do you use in preparing food in a week? Mean=128.08 sd=73.00	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y
P048: How much did you spend last week on alcohol? (mean=2.054, sd=10.845)	N	N	N	N	N	N	N	Y	N	N	N	Y	N	N	N	N
P048: How much did you spend last week on soft drinks? (mean=10.14, sd=15.02)	N	N	Y	Y	N	Y	Y	Y	N	N	Y	Y	N	N	Y	Y
P05003: How much did you spend in last 6 months on clothes for girls mean=50.80 sd=100.16	N	N	Y	Y	N	N	Y	Y	N	N	Y	Y	N	N	Y	Y
P048: How much did you spend last week on medicine? (mean=42.83, sd=13.46)	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

n.b. Based on control and treatment “poor” households in Table 1.

Table 7, Part 3 (Children's Health Module, Individual Level Variables)
Whether Different Samples Satisfy Required Sample Size for Two-sided Test of Equality of Proportions or Equality of Means
significance level =0.05, power =0.90

Variable of Interest	Quarter of Initial Control Sample and Equal Treatment Phase-in Sample (n2/n1) ratio=1.0 treatment impact magnitude				Initial Evaluation Sample (n2/n1) ratio= 0.6204 treatment impact magnitude				Half of Initial Control Sample and Initial Treatment Sample (n2/n1) ratio = 0.3102 treatment impact magnitude				Quarter of Initial Control Sample and Initial Treatment Sample (n2/n1) ratio = 0.1551 treatment impact magnitude			
	5%	10%	25%	50%	5%	10%	25%	50%	5%	10%	25%	50%	5%	10%	25%	50%
P056: Was child sick in the last 4 weeks? (p=0.2955)	N	N	Y (4 yrs)	Y	N	Y	Y	Y	N	N	Y (2 yrs)	Y	N	N	Y	Y
P057: Duration of sickness (mean=7.0857 sd=5.1458)	N	N	Y (4 yrs)	Y	N	Y	Y	Y	N	N	Y (2 yrs)	Y	N	N	Y	Y
P060: Saw a health helper, doctor, nurse or dentist regarding illness (p=0.6593)	N	N	Y (3 yrs)	Y	N	Y	Y	Y	N	N	Y (2 yrs)	Y	N	N	Y	Y
P068: Is child less than one year old currently being breast fed? (p=0.7492)	N	N	Y	*	N	Y	Y	*	Y (5 yrs)	Y (2 yrs)	Y	*	N	N	Y	*
P069: Did you take child to health center to determine weight and height? (p=0.8027)	N	Y (3 yrs)	*	*	Y (3 yrs)	Y	*	*	Y (4 yrs)	Y (2 yrs)	*	*	Y (6 yrs)	Y (2 yrs)	*	*
P073-P076: Does child have TB, triple vaccine, polio and measles vaccines? (p=0.8641)	Y (6 yrs)	Y (2 yrs)	*	*	Y (2 yrs)	Y	*	*	Y (3 yrs)	Y	*	*	Y (4 yrs)	Y	*	*

N.B. Based on Table2, using youngest ages in 0-6 age range necessary to get required sample sizes that are indicated in Table 4 (i.e., "2 yrs" means that ages 0 and 1 must be combined, "3 yrs" means that ages 0- must be combined). "x" means that the treatment impact would cause the proportion to exceed 1.

Table 7, Part 4 (Household and Women's Health Module)
Whether Different Samples Satisfy Required Sample Size for Two-sided Test of Equality of Proportions or Equality of Means
significance level =0.05, power =0.90

Variable of Interest	Quarter of Initial Control Sample and Equal Treatment Phase-in Sample				Initial Evaluation Sample				Half of Initial Control Sample and Initial Treatment Sample				Quarter of Initial Control Sample and Initial Treatment Sample			
	(n2/n1) ratio=1.0				(n2/n1) ratio= 0.6204				(n2/n1) ratio = 0.3102				(n2/n1) ratio = 0.1551			
	treatment impact magnitude				treatment impact magnitude				treatment impact magnitude				treatment impact magnitude			
	5%	10%	25%	50%	5%	10%	25%	50%	5%	10%	25%	50%	5%	10%	25%	50%
P077: Do you know how to prepare oral serum? (p=0.8223)	Y	Y	*	*	Y	Y	*	*	Y	Y	*	*	Y	Y	*	*
P078: Do you boil or put bleach in the water? (p=0.7864)	N	Y	Y	*	Y	Y	Y	*	Y	Y	Y	*	N	Y	Y	*
P082*[(1)]: Did you or your family go to a medical center to see a doctor in the last year? (p=0.8084)	Y	Y	*	*	Y	Y	*	*	Y	Y	*	*	Y	Y	*	*
P102: Did you see a doctor or nurse during this pregnancy? (p=0.5776)	N	Y	Y	Y	N	Y	Y	Y	N	Y	Y	Y	N	Y	Y	Y
P120-P122*[(2)]: Are you currently or have you ever used contraceptives or are you sterilized? (p=0.4168)	N	N	Y	Y	N	Y	Y	Y	N	Y	Y	Y	N	N	Y	Y

n.b. Based on control and treatment "poor" households in Table 1.

Table 7, Part 5 (Women's Status Module)
Whether Different Samples Satisfy Required Sample Size for Two-sided Test of Equality of Proportions or Equality of Means
significance level =0.05, power =0.90

Variable of Interest	**Quarter of Initial Control Sample and Equal Treatment Phase-in Sample** (n2/n1) ratio=1.0 treatment impact magnitude				**Initial Evaluation Sample** (n2/n1) ratio= 0.6204 treatment impact magnitude				**Half of Initial Control Sample and Initial Treatment Sample** (n2/n1) ratio = 0.3102 treatment impact magnitude				**Quarter of Initial Control Sample and Initial Treatment Sample** (n2/n1) ratio = 0.1551 treatment impact magnitude			
	5%	10%	25%	50%	5%	10%	25%	50%	5%	10%	25%	50%	5%	10%	25%	50%
P129: If the woman has extra income, she should decide what to do. (p=0.1947)	N	N	Y	Y	N	N	Y	Y	N	N	Y	Y	N	N	Y	Y
P137: Women must always obey men (p=0.794)	N	Y	Y	*	Y	Y	Y	*	Y	Y	Y	*	Y	Y	Y	*
P137: Women may work out of the household (p=0.708)	N	Y	Y	*	Y	Y	Y	*	Y	Y	Y	*	N	Y	Y	*
P137: Women and men have the same rights (p=0.883)	Y	Y	*	*	Y	Y	*	*	Y	Y	*	*	Y	Y	*	*

n.b. Based on control and treatment "poor" households in Table 1. "*" means that the treatment impact would cause the proportion to exceed 1.

Table 8, Part 1 (Education Module, Individual Level Variables)
Required Sample Size for Two-sided Test of Equality of Proportions or Equality of Means
significance level α=0.05, power β=0.90
For Different Variables of Interest and Varying Treatment/Control Proportions[1]

	(n2/n1) ratio=1.0 (half initial control group and equal size treatment group) treatment impact magnitude				(n2/n1) ratio= 0.6204 (half initial treatment group, half initial control group) treatment impact magnitude				(n2/n1) ratio = 0.4138 (3/4 initial treatment group, half initial control group) treatment impact magnitude				(n2/n1) ratio = 0.0.3102 (initial treatment group, half initial control group) treatment impact magnitude			
	5%	10%	25%	50%	5%	10%	25%	50%	5%	10%	25%	50%	5%	10%	25%	50%
Variable of Interest																
P006: proportion attending school	1418	318	*	*	1862	420	*	*	2446	554	*	*	3030	687	*	*
(p=0.8444)	1418	318			1156	261			1013	230			940	214		
P007*[2]: not in school because of resource constraints	8328	2096	338	83	10876	2737	442	109	14227	3581	579	142	17588	4426	715	175
(p=0.5044)	8328	2096	338	83	6748	1699	275	68	5888	1482	240	59	5456	1373	222	55
P008* and P012*[3]: imputed years of education	3799	950	152	38	4961	1241	199	50	6490	1623	260	65	8023	2006	321	81
(mean=4.244, sd=2.853)	3799	950	152	38	3078	770	124	32	2686	672	108	27	2489	623	100	26
P009: how is child doing at school? (good=1)	5325	1323	204	45	6959	1731	268	59	9108	2266	351	78	11262	2803	435	96
(p = 0.6114)	5325	1323	204	45	4318	1074	167	37	3769	938	146	33	3494	870	135	30
P011: expected grade completion level																
no more than primary	41202	10552	1805	497	53756	13754	2348	645	70267	17966	3060	837	86286	22189	3774	1030
(p=0.173)	41202	10552	1805	497	33351	8533	1457	401	29077	7435	1267	347	26934	6884	1171	320
secondary but not higher	8103	2056	332	81	10674	2685	433	106	13963	3513	567	139	17262	4342	700	171
(p=0.509)	8103	2056	332	81	6623	1666	269	66	5778	1454	235	58	5355	1347	218	54
more than secondary	18387	4683	787	210	23997	6108	1025	273	31375	7982	1338	355	38775	9862	1651	438
(p=0.318)	18387	4683	787	210	14888	3790	636	170	12983	3303	554	147	12029	3060	513	136
P014: number of days absent in last 4 weeks	397	330	206	113	518	430	269	148	677	563	351	193	837	695	434	239
(mean=0.4691, sd=1.9588)	397	330	206	113	322	267	167	92	281	233	146	80	260	216	135	75

(1) The treatment impacts considered are always in the direction of increasing the mean or proportion. A * appears when the treatment impact would cause the proportion to exceed 1.
(2) Variable constructed from P007. Equals 1 if P007 equals 01, 02, or 03
(3) Variable constructed from P004, P006, P008, and P012. For individuals currently in school, construct Age-6-P012. For individuals not currently in school, construct Age-6-P008 if P008<8, else 0. Recode negative values for this group to 0.

Table 8, Part 2 (Education and Consumption Module, Household Level Variables)
Required Sample Size for Two-sided Test of Equality of Proportions or Equality of Means
significance level α=0.05, power β=0.90
For Different Variables of Interest and Varying Treatment/Control Proportions[(1)]

Variable of Interest	(n2/n1) ratio=1.0 (half initial control group and equal size treatment group) treatment impact magnitude				(n2/n1) ratio= 0.6204 (half initial treatment group, half initial control group) treatment impact magnitude				(n2/n1) ratio = 0.4138 (3/4 initial treatment group, half initial control group) treatment impact magnitude				(n2/n1) ratio = 0.0.3102 (initial treatment group, half initial control group) treatment impact magnitude			
	5%	10%	25%	50%	5%	10%	25%	50%	5%	10%	25%	50%	5%	10%	25%	50%
P025: Are you or your husband in the school parents association? (p=0.3389)	16708 16708	4251 4251	712 712	189 189	21808 13530	5546 3441	928 576	245 152	28514 11800	7248 3000	1211 502	320 133	35239 10932	8955 2778	1495 464	394 123
P038: How much do you spend on fruits and vegetables? (mean=35.29 sd=31.19)	6567 6567	1642 1642	263 263	66 66	8575 5320	2144 1331	343 213	86 54	11218 4643	2805 1161	449 186	113 47	13867 4302	3467 1076	555 173	139 44
P035b01: Home-produced chicken? (for example) (p=0.6641)	4201 4201	1034 1034	154 154	31 31	5493 3408	1354 841	203 126	41 26	7192 2977	1774 735	266 111	53 22	8896 2760	2195 681	329 103	64 20
P038: How much do you spend on food of animal origin? (mean=49.31 sd=40.78)	5750 5750	4069 4069	230 230	58 58	7509 4659	5313 3297	301 187	76 48	9822 4065	6950 2876	393 163	99 41	12142 3767	8592 2666	486 151	122 38
P046: How much money do you use in preparing food in a week? (mean=128.08 sd=73.00)	2731 2731	683 683	110 110	28 28	3567 2213	892 554	143 89	36 23	4665 1931	1167 483	187 78	47 20	5767 1789	1442 448	231 72	58 18
P048: How much did you spend last week on alcohol? (mean=2.054, sd=10.845)	234339 234339	58585 58585	9374 9374	2344 2344	306031 189862	76508 47466	12242 7595	3061 1900	400325 165655	100082 41414	16013 6627	4004 1657	494892 153516	123723 38379	19796 6141	4949 1536
P048: How much did you spend last week on soft drinks? (mean=10.14, sd=15.02)	18444 18444	4611 4611	738 738	185 185	24087 14944	6022 3737	964 599	241 150	31508 13039	7877 3260	1261 522	316 131	38951 12083	9738 3021	1559 484	390 121

(1) The treatment impacts considered are always in the direction of increasing the mean or proportion. A * appears when the treatment impact would cause the proportion to exceed 1.

Table 8, Part 2 continued (Education and Consumption Module, Household Level Variables)
Required Sample Size for Two-sided Test of Equality of Proportions or Equality of Means
significance level α=0.05, power β=0.90
For Different Variables of Interest and Varying Treatment/Control Proportions[1]

Variable of Interest	(n2/n1) ratio=1.0 (half initial control group and equal size treatment group) treatment impact magnitude				(n2/n1) ratio= 0.6204 (half initial treatment group, half initial control group) treatment impact magnitude				(n2/n1) ratio = 0.4138 (3/4 initial treatment group, half initial control group) treatment impact magnitude				(n2/n1) ratio = 0.0.3102 (initial treatment group, half initial control group) treatment impact magnitude			
	5%	10%	25%	50%	5%	10%	25%	50%	5%	10%	25%	50%	5%	10%	25%	50%
P05003: How much did you spend in last 6 months on clothes for girls mean=50.80 sd=100.16	32678 32678	8170 8170	1308 1308	327 327	42675 26476	10669 6620	1707 1060	427 265	55824 23100	13956 5775	2233 925	559 232	69011 21408	17253 5352	2761 857	691 215
P048: How much did you spend last week on medicine? (mean=42.83, sd=13.46)	831 831	208 208	34 34	9 9	1085 674	272 169	44 28	11 7	1419 588	355 147	57 24	15 7	1754 545	439 137	71 23	18 6

(1) The treatment impacts considered are always in the direction of increasing the mean or proportion. A * appears when the treatment impact would cause the proportion to exceed 1.

Table 8, Part 3 (Children's Health Module, Individual Level Variables)
Required Sample Size for Two-sided Test of Equality of Proportions or Equality of Means
significance level α=0.05, power β=0.90
For Different Variables of Interest and Varying Treatment/Control Proportions[1]

	(n2/n1) ratio=1.0 (half initial control group and equal size treatment group) treatment impact magnitude				(n2/n1) ratio= 0.6204 (half initial treatment group, half initial control group) treatment impact magnitude				(n2/n1) ratio = 0.4138 (3/4 initial treatment group, half initial control group) treatment impact magnitude				(n2/n1) ratio = 0.0.3102 (initial treatment group, half initial control group) treatment impact magnitude			
Variable of Interest	5%	10%	25%	50%	5%	10%	25%	50%	5%	10%	25%	50%	5%	10%	25%	50%
P056: Was child sick in the last 4 weeks? (p=0.2955)	20391 20391	5199 5199	879 879	236 236	26611 16510	6780 4207	1145 711	306 190	34792 14397	8860 3667	1494 619	399 166	42996 13338	10945 3396	1843 572	491 153
P057: Duration of sickness (mean=7.0857 sd=5.1458)	4433 4433	1109 1109	178 178	45 45	5790 3593	1448 899	232 144	58 36	7573 3134	1894 784	303 126	76 32	9362 2905	2341 727	375 117	94 30
P060*[2]: Saw a health helper, doctor, nurse or dentist regarding illness (p=0.6593)	4287 4287	1059 1059	159 159	32 32	5605 3478	1387 861	208 130	42 27	7338 3037	1817 752	274 114	55 23	9076 2816	2249 698	338 105	67 21
P068*[3] Is child less than one year old currently being breast fed? (p=0.7492)	2715 2715	654 654	88 88	*	3554 2205	858 533	117 73	*	4658 1928	1126 466	153 64	*	5764 1788	1395 433	189 59	*
P069: Did you take child to health center to determine weight and height? (p=0.8027)	1954 1954	455 455	*	*	2562 1590	599 372	*	*	3360 1391	787 326	*	*	4161 1291	976 303	*	*
P073-P076: Does child have TB, triple vaccine, polio and measles vaccines? (p=0.8641)	1184 1184	258 258	*	*	1558 967	342 213	*	*	2047 848	451 187	149 62	46 20	2538 788	559 174	*	*

(1) The treatment impacts considered are always in the direction of increasing the mean or proportion. A * appears when the treatment impact would cause the proportion to exceed 1.
(2) Variable constructed from P060.
(3) This variable constructed from P068, restricting population to children less than one year old.

Table 8, Part 4 (Household and Women's Health Module)
Required Sample Size for Two-sided Test of Equality of Proportions or Equality of Means
significance level α=0.05, power β=0.90
For Different Variables of Interest and Varying Treatment/Control Proportions

	(n2/n1) ratio=1.0 (half initial control group and equal size treatment group) treatment impact magnitude				(n2/n1) ratio= 0.6204 (half initial treatment group, half initial control group) treatment impact magnitude				(n2/n1) ratio = 0.4138 (3/4 initial treatment group, half initial control group) treatment impact magnitude				(n2/n1) ratio = 0.0.3102 (initial treatment group, half initial control group) treatment impact magnitude			
Variable of Interest	5%	10%	25%	50%	5%	10%	25%	50%	5%	10%	25%	50%	5%	10%	25%	50%
P077: Do you know how to prepare oral serum? (p=0.8223)	1695 1695	389 389	108 108	32 32	2223 1380	513 319	140 87	42 27	2918 1208	675 280	180 75	53 22	3614 1122	838 260	220 69	65 21
P078: Do you boil or put bleach in the water? (p=0.7864)	2176 2176	513 513	64 64	36 36	2852 1770	675 419	84 53	47 30	3740 1548	887 368	111 46	60 25	4630 1437	1099 341	136 43	73 23
P082*[(2)]: Did you or your family go to a medical center to see a doctor in the last year? (p=0.8084)	1876 1876	436 436	115 115	34 34	2460 1527	574 357	148 92	44 28	3228 1336	755 313	191 80	56 24	3997 1240	936 291	234 73	68 22
P102: Did you see a doctor or nurse during this pregnancy? (p=0.5776)	6145 6145	1534 1534	241 241	56 56	8028 4981	2005 1244	316 197	73 46	10505 4347	2625 1087	414 172	96 40	12989 4030	3246 1007	512 159	118 37
P120-P122*[(3)]: Are you currently or have you ever used contraceptives or are you sterilized? (p=0.4168)	11980 11980	3021 3021	499 499	129 129	15640 9704	3942 2446	651 404	168 105	20454 8464	5154 2133	851 353	219 91	25281 7843	6369 1976	1051 327	270 84

(1) Variable constructed from P082 by recoding 1,2, or 3 as went to a doctor.
(2) Constructed variable from P120, P121, P122

Table 8, Part 5 (Women's Status Module)
Required Sample Size for Two-sided Test of Equality of Proportions or Equality of Means
significance level α=0.05, power β=0.90
For Different Variables of Interest and Varying Treatment/Control Proportions[(1)]

	(n2/n1) ratio=1.0 (half initial control group and equal size treatment group) treatment impact magnitude				(n2/n1) ratio= 0.6204 (half initial treatment group, half initial control group) treatment impact magnitude				(n2/n1) ratio = 0.4138 (3/4 initial treatment group, half initial control group) treatment impact magnitude				(n2/n1) ratio = 0.0.3102 (initial treatment group, half initial control group) treatment impact magnitude			
Variable of Interest	5%	10%	25%	50%	5%	10%	25%	50%	5%	10%	25%	50%	5%	10%	25%	50%
P129*[(1)]: If the woman has extra income, she should decide what to do with it. (p=0.1947)	35880 35880	9090 9090	1555 1555	427 427	46814 29044	11850 7352	2022 1255	553 344	61195 25323	15479 6406	2637 1092	719 298	75617 23457	19118 5931	3252 1009	885 275
P137: Women must always obey men (p=0.794)	2068 2068	485 485	59 59	*	2711 1682	639 397	78 49	*	3556 1472	840 348	102 43	58 25	4403 1366	1041 323	125 39	*
P137: Women may work out of the household (p=0.708)	3393 3393	826 826	118 118	*	4439 2754	1083 672	156 97	*	5814 2406	1420 588	205 85	76 32	7193 2232	1758 546	253 79	*
P137: Women and men have the same rights (p=0.883)	967 967	202 202	*	*	1274 791	269 167	*	*	1676 694	355 147	136 57	43 18	2079 645	440 137	*	*

(1) The treatment impacts considered are always in the direction of increasing the mean or proportion. A * appears when the treatment impact would cause the proportion to exceed 1.

Table 9, Part 1 (Education Module, Individual Level Variables)
Whether Different Samples Satisfy Required Sample Size for Two-sided Test of Equality of Proportions or Equality of Means
significance level α=0.05, power β=0.90

Variable of Interest	Half Initial Control Sample and Equal-Size Treatment Sample (n2/n1) ratio=1.0				Half Initial Treatment Group and Half Initial Evaluation Sample (n2/n1) ratio= 0.6204				Half of Initial Control Sample and ¾ Initial Treatment Sample (n2/n1) ratio = 0.4138				Half of Initial Control Sample and Initial Treatment Sample (n2/n1) ratio = 0.3102			
	treatment impact magnitude				treatment impact magnitude				treatment impact magnitude				treatment impact magnitude			
	5%	10%	25%	50%	5%	10%	25%	50%	5%	10%	25%	50%	5%	10%	25%	50%
P006: proportion attending school (p=0.8446)	Y (both, 4 yrs)	Y (both, 2 yrs)	*	*	Y (both, 4 yrs)	Y (both, 2 yrs)	*	*	Y (both, 4 yrs)	Y (both, 2 yrs)	*	*	Y (both, 4 yrs)	Y (both, 2 yrs)	*	*
P007: not in school because of resource constraints (p=0.5044)	N	N	Y (both, 3yrs)	Y (both)	N	N	Y (both, 2yrs)	Y (both)	N	N	Y (both, 2yrs)	Y	N	N	Y (both, 2yrs)	Y
P008 and P012: imputed years of education (mean=4.244, sd=2.853)	Y (both, 10 yrs)	Y (both, 4 yrs)	Y (both)	Y	Y (both, 9 yrs)	Y (both, 3 yrs)	Y (both)	Y	Y (both, 9 yrs)	Y (both, 3 yrs)	Y (both)	Y	Y (both, 9 yrs)	Y (both, 3 yrs)	Y (both)	Y
P009: how is child doing at school? (good=1) (p = 0.6114)	N	Y (both, 6 yrs)	Y (both, 2 yrs)	Y (both)	N	Y (both, 5 yrs)	Y (both, 2yrs)	Y (both)	N	Y (both, 5 yrs)	Y (both, 2 yrs)	Y	Y (both, 11 yrs)	Y (both, 5 yrs)	Y (both, 2 yrs)	Y
P011: expected grade completion level																
no more than primary (p=0.173)	N	N	Y (both, 7 yrs)	Y (both, 4 yrs)	N	N	Y (both, 6 yrs)	Y (both, 3 yrs)	N	N	Y (both, 6 yrs)	Y (both, 3yrs)	N	N	Y (both, 6 yrs)	Y (both, 3 yrs)
secondary but not higher (p=0.509)	N	Y (both, 8 yrs)	Y (both, 3 yrs)	Y (both)	N	Y (both, 7 yrs)	Y (both, 6yrs)	Y (both, 3 yrs)	N	Y (both, 7 yrs)	Y (both, 2 yrs)	Y (both)	N	Y (both, 6 yrs)	Y (both, 2 yrs)	Y (both)
more than secondary (p=0.318)	N	N	Y (both, 5 yrs)	Y (both, 2 yrs)	N	N	Y (both, 4 yrs)	Y (both, 2 yrs)	N	Y (both, 7yrs)	Y (both, 2 yrs)	Y (both, 2 yrs)	N	Y (both, 10 yrs)	Y (both, 4 yrs)	Y (both, 2 yrs)
P014: number of days absent in last 4 weeks (mean=0.4691, sd=1.9588)	Y (both, 3 yrs)	Y (both, 3 yrs)	Y (both, 2 yrs)	Y (both, 2 yrs)	Y (both, 3 yrs)	Y (both, 3 yrs)	Y (both, 2 yrs)	Y (both, 2 yrs)	Y (both, 3 yrs)	Y (both, 2 yrs)	y (both, 2 yrs)	Y (both)	Y (both, 3 yrs)	Y (both, 2 yrs)	Y (both, 2 yrs)	Y (both)

N.B. Based on Table 3, using oldest ages in 6-16 age range necessary to get required sample sizes that are indicated in Table 4 (i.e., "2 yrs" means that ages 15 and 16 must be combined, "3 yrs" means that ages 14-16 must be combined). "Both" means that both females and males must be combined or the years doubled. "*" means that the treatment impact would cause the proportion to exceed 1.

Table 9, Part 2 (Education and Consumption Module, Household Level Variables)
Whether Different Samples Satisfy Required Sample Size for Two-sided Test of Equality of Proportions or Equality of Means
significance level α=0.05, power β=0.90

Variable of Interest	Half Initial Control Sample and Equal-Size Treatment Sample (n2/n1) ratio=1.0				Half Initial Treatment Group and Half Initial Evaluation Sample (n2/n1) ratio= 0.6204				Half of Initial Control Sample and ¾ Initial Treatment Sample (n2/n1) ratio = 0.4138				Half of Initial Control Sample and Initial Treatment Sample (n2/n1) ratio = 0.3102			
	treatment impact magnitude				treatment impact magnitude				treatment impact magnitude				treatment impact magnitude			
	5%	10%	25%	50%	5%	10%	25%	50%	5%	10%	25%	50%	5%	10%	25%	50%
P025: Are you or your husband in the school parents association? p=0.3389	N	N	Y	Y	N	N	Y	Y	N	N	Y	Y	N	N	Y	Y
P038: How much do you spend on fruits and vegetables? mean=35.29 sd=31.19	N	Y	Y	Y	N	Y	Y	Y	N	Y	Y	Y	N	Y	Y	Y
P035b01: Home-produced chicken? p=0.6641	N	Y	Y	Y	N	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Y
P038: How much do you spend on food of animal origin? Mean=49.31 sd=40.78	N	N	Y	Y	N	N	Y	Y	N	N	Y	Y	N	N	Y	Y
P046: How much money do you use in preparing food in a week? Mean=128.08 sd=73.00	N	Y	Y	Y	N	Y	Y	Y	N	Y	Y	Y	N	Y	Y	Y
P048: How much did you spend last week on alcohol? (mean=2.054, sd=10.845)	N	N	N	N	N	N	N	Y	N	N	N	Y	N	N	N	Y
P048: How much did you spend last week on soft drinks? (mean=10.14, sd=15.02)	N	N	Y	Y	N	N	Y	Y	N	N	Y	Y	N	N	Y	Y
P05003: How much did you spend in last 6 months on clothes for girls mean=50.80 sd=100.16	N	N	Y	Y	N	N	Y	Y	N	N	Y	Y	N	N	Y	Y
P048: How much did you spend last week on medicine? (mean=42.83, sd=13.46)	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

n.b. Based on control and treatment “poor” households in Table 1.

Table 9, Part 3 (Children's Health Module, Individual Level Variables)
Whether Different Samples Satisfy Required Sample Size for Two-sided Test of Equality of Proportions or Equality of Means
significance level α=0.05, power β=0.90

Variable of Interest	**Half Initial Control Sample and Equal-Size Treatment Sample** (n2/n1) ratio=1.0				**Half Initial Control Sample and Half Initial Treatment Sample** (n2/n1) ratio= 0.6204				**Half of Initial Control Sample and ¾ Initial Treatment Sample** (n2/n1) ratio = 0.4138				**Half of Initial Control Sample and Initial Treatment Sample** (n2/n1) ratio = 0.3102			
	treatment impact magnitude				treatment impact magnitude				treatment impact magnitude				treatment impact magnitude			
	5%	10%	25%	50%	5%	10%	25%	50%	5%	10%	25%	50%	5%	10%	25%	50%
P056: Was child sick in the last 4 weeks? (p=0.2955)	N	N	Y (3 yrs)	Y	N	N	Y (3 yrs)	Y	N	N	Y (2 yrs)	Y	N	N	Y (2 yrs)	Y
P057: Duration of sickness (mean=7.0857 sd=5.1458)	N	N	Y (2 yrs)	Y	N	N	Y (2 yrs)	Y	N	N	Y (2 yrs)	Y	N	N	Y (2 yrs)	Y
P060: Saw a health helper, doctor, nurse or dentist regarding illness (p=0.6593)	N	N	Y (2 yrs)	Y	N	N	Y (2 yrs)	Y	N	N	Y (2 yrs)	Y	N	N	Y (2 yrs)	Y
P068: Is child less than one year old currently being breast fed? (p=0.7492)	N	Y (2 yrs)	Y	*	Y (6 yrs)	Y (2 yrs)	Y	*	Y (5 yrs)	Y (2 yrs)	Y	*	Y (5 yrs)	Y (2 yrs)	Y	*
P069: Did you take child to health center to determine weight and height? (p=0.8027)	Y (5 yrs)	Y (2 yrs)	*	*	Y (5 yrs)	Y (2 yrs)	*	*	Y (4 yrs)	Y (2 yrs)	*	*	Y (4 yrs)	Y	*	*
P073-P076: Does child have TB, triple vaccine, polio and measles vaccines? (p=0.8641)	Y (4 yrs)	Y	*	*	Y (3 yrs)	Y	*	*	Y (3 yrs)	Y	*	*	Y (3 yrs)	Y	*	*

N.B. Based on Table2, using youngest ages in 0-6 age range necessary to get required sample sizes that are indicated in Table 4 (i.e., "2 yrs" means that ages 0 and 1 must be combined, "3 yrs" means that ages 0- must be combined). "*" means that the treatment impact would cause the proportion to exceed 1.

Table 9, Part 4 (Household and Women's Health Module)
Whether Different Samples Satisfy Required Sample Size for Two-sided Test of Equality of Proportions or Equality of Means
significance level α=0.05, power β=0.90

Variable of Interest	**Half Initial Control Sample and Equal-Size Treatment Sample** (n2/n1) ratio=1.0				**Half Initial Control Sample and Half Initial Treatment Sample** (n2/n1) ratio= 0.6204				**Half of Initial Control Sample and ¾ Initial Treatment Sample** (n2/n1) ratio = 0.4138				**Half of Initial Control Sample and Initial Treatment Sample** (n2/n1) ratio = 0.3102			
	treatment impact magnitude				treatment impact magnitude				treatment impact magnitude				treatment impact magnitude			
	5%	10%	25%	50%	5%	10%	25%	50%	5%	10%	25%	50%	5%	10%	25%	50%
P077: Do you know how to prepare oral serum? (p=0.8223)	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
P078: Do you boil or put bleach in the water? (p=0.7864)	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
P082*[(1)]: Did you or your family go to a medical center to see a doctor in the last year? (p=0.8084)	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
P102: Did you see a doctor or nurse during this pregnancy? (p=0.5776)	N	Y	Y	Y	N	Y	Y	Y	N	Y	Y	Y	N	Y	Y	Y
P120-P122*[(2)]: Are you currently or have you ever used contraceptives or are you sterilized? (p=0.4168)	N	N	Y	Y	N	N	Y	Y	N	Y	Y	Y	N	Y	Y	Y

n.b. Based on control and treatment "poor" households in Table 1.

Table 9, Part 5 (Women's Status Module)
Required Sample Size for Two-sided Test of Equality of Proportions or Equality of Means
significance level α=0.05, power β=0.90
For Different Variables of Interest and Varying Treatment/Control Proportions[(1)]

	(n2/n1) ratio=1.0 (half initial control group and equal size treatment group) treatment impact magnitude				(n2/n1) ratio= 0.6204 (half initial treatment group, half initial control group) treatment impact magnitude				(n2/n1) ratio = 0.4138 (3/4 initial treatment group, half initial control group) treatment impact magnitude				(n2/n1) ratio = 0.0.3102 (initial treatment group, half initial control group) treatment impact magnitude			
Variable of Interest	5%	10%	25%	50%	5%	10%	25%	50%	5%	10%	25%	50%	5%	10%	25%	50%
P129*[(1)]: If the woman has extra income, she should decide what to do with it. (p=0.1947)	N	N	Y	Y	N	N	Y	Y	N	N	Y	Y	N	N	Y	Y
P137: Women must always obey men (p=0.794)	Y	Y	Y	*	Y	Y	Y	*	Y	Y	Y	*	Y	Y	Y	*
P137: Women may work out of the household (p=0.708)	N	Y	Y	*	N	Y	Y	*	N	Y	Y	*	Y	Y	Y	*
P137: Women and men have the same rights (p=0.883)	Y	Y	*	*	Y	Y	*	*	Y	Y	*	*	Y	Y	*	*

(1) The treatment impacts considered are always in the direction of increasing the mean or proportion. A * appears when the treatment impact would cause the proportion to exceed 1.