Proposal

To Conduct

Policy Impact Evaluations on Human Capital Investment Policies and Programs

Assessing the Impact of Ishraq Intervention

A Second-Chance Program for Out-of-school Rural Adolescent Girls in Egypt

Submitted to : PEP-AUSAID Policy Impact Evaluation Research Initiative (PIERI)

By : Population Council, West Asia and North Africa Regional Office

Contact Person: Ragui Assaad, Regional Director

Date : May 23, 2008
Abstract:

Egypt faces continuing and sometimes widening disparities between its urban and rural populations in key health, social, and economic indicators. Development programs often miss a disadvantaged population group caught between childhood and adulthood— rural adolescent girls. Ishraq is an intervention designed to address this problem. It is a non-formal second-chance education and social support program aiming at promoting a healthful and active transition to adulthood for disadvantaged out-of-school rural girls. It prepares them to make informed decisions about life issues such as schooling, marriage, and livelihood opportunities.

This proposed project aims at carrying out a program impact evaluation of the Ishraq project using a quasi-experimental design, including baseline and end-line surveys in randomly selected control and intervention villages. The surveys will be administered to a panel of out-of-school adolescent girls who are 12-15 years old. The impact of the program is assessed by comparing end-of-program responses to baseline responses across participants (by level of participation), non-participants in intervention villages, and girls in control villages. Over the next three years, Ishraq will be implemented in as many as 30 villages reaching about 1500 girls in three governorates in Upper Egypt.

A. Aims and Objectives:

a. Study Overview:

The benefits of monitoring, evaluation and impact assessment have been recognized by the partners implementing the Ishraq program from the start. A monitoring and evaluation component, which made use of both qualitative and quantitative methods was in fact incorporated into the pilot phase of the program, which was carried out from 2001 to 2004. The results of these evaluation efforts are reported in Brady et al. (2007). Because the pilot program involved only 277 participants, it did not allow for a very detailed assessment of the project’s impact. Moreover, the impact of a small pilot is often difficult to replicate at scale. We are therefore proposing to conduct a rigorous impact evaluation during the expansion phase of the project which will involve at least 1500 project participants in about 30 villages. Due to budgetary constraints, activities included in this proposal will cover data collection in five villages. Other funding sources are currently being sought to cover the cost of collecting data from all intervention and control villages.

b. Main Research Questions and Core Research Objectives

The core research objective is to assess the impact of a 20-month non-formal education intervention on the program participants and the communities in which they live. This assessment would serve to inform the continued implementation of the intervention and its expansion, as well as inform the design and implementation of other efforts to reach one of the most vulnerable and disadvantaged target groups in Egyptian society, namely rural out-of-school adolescent girls. One of the core philosophies of the Ishraq program is that in order to reach girls in these socially-conservative settings, we have to first break their social isolation and create safe public spaces in which they can gather, make friendships, learn, and play. Thus a


2
main objective of the study is to gauge the extent to which such an approach actually manages to alter the course of an adolescent girl’s life by changing her knowledge skills and attitudes, as well as changing the attitude of those around her.

The specific areas in which we are going to assess the impact of the projects are: (i) functional literacy, cognitive skills and continuing schooling, (ii) attitudes about marriage and childbearing, (iii) knowledge about nutrition, hygiene, and reproductive health, (iv) attitudes about harmful traditional practices, such as female genital mutilation/cutting, (v) social isolation, peer networks, and participation in group or community activities, and (vi) gender norms, gender-based violence, mobility, and agency, including attitudes about participation in sports and recreation activities. The study will also strive to measure parents’ and brothers’ attitudes about these issues. A review of the literature on the impact of those short-term outcomes on longer term outcomes as these girls transition to adulthood will also be conducted.

While the study period is not long enough to gauge the impact of the program on longer term outcomes such as ultimate educational attainment, age at marriage, fertility, child health and education, participation in economic activity, and empowerment in household decision-making, it will establish a database that can be used to undertake such studies at a later date by tracking respondents at a later date.

B. Background and Policy Relevance:

In this section, we review the literature on program evaluation. This section is primarily informed by Heckman, Lalonde & Smith (1999), Blundell & Costa Dias (2002), and Smith (2004). We cover the main issues as well as the new developments of that literature with an emphasis on challenges facing the estimation of treatment impacts. As highlighted in the next section, these primarily include selection bias resulting from the voluntary nature of participation in the program.

The importance of program evaluation has significantly increased over the last few decades, given the expansion of social programs that had started in the US since the 1960s and expanded in the US and in various other countries and the growing need to assess their impact. Programs geared towards employment and training in particular have received focus. The last two decades witnessed major technical advancements in evaluation approaches which were aided by concurrent econometric advancements such as those in the areas of non-parametric and semi-parametric estimation.

In order to evaluate the impact of a program, we need to know the difference between the outcomes for participants if they participated in the program and the outcomes had they not participated, where the latter is not observed. Accordingly, the primary problem in identifying treatment effects becomes the estimation of that counterfactual. This explains why statisticians tend to think of the evaluation estimation as a missing data problem.

The counterfactual can be proxied by the average outcomes of untreated units (for example from a comparison or control group). This is not a problem in the case of random social experiments where assignment into treatment is random. However, in many experiments,

---

2A review of the different treatment parameters used in the literature is provided in many papers (for example, see Heckman & Lalonde & Smith 1999, Blundell & Costa Dias 2002, and Smith 2004).
participation (within the intervention group) is voluntary, thereby violating randomness. Additionally, in many cases, random trials are not feasible, and are usually expensive and hard to implement thus calling for the use of non-experimental (econometric) methods where a comparison group with “comparable” characteristics is constructed. As a result, a selection bias problem frequently arises when the (to-be) treated and the untreated differ in (often unobservable) characteristics, which could potentially result in mixing-up the treatment effect with pre-program differences. This is why econometricians tend to think of evaluation estimation as a selection problem, hence the close relationship between evaluation research and the Heckman selection literature.

Lalonde (1986) produced an influential paper where he compared non-experimental econometric estimates (where the comparison group was constructed based on population surveys) with a random experiment estimate of the same training program. As results of the alternative techniques were different and many were not close to the experimental estimates, trust in these techniques’ reliability was shaken.

This paper triggered a significant response and led to the production of a series of papers that use a random experiment as a benchmark for assessing different econometric methods. Of these, some re-analyzed the same dataset Lalonde used (e.g., Heckman & Hotz 1989, Dehejia & Wahba 1999 and Smith & Todd 2004). For example, Heckman and Hotz (1989) concluded that using specification tests requiring pre- and post- program data can rule out inappropriate estimators. Other papers analyzed the Job Training Partnership Act (JTPA) (e.g., Heckman & Ichimura & Todd 1997 and Heckman & Ichimura & Smith & Todd 1998). In general, these papers helped explain situations in which particular econometric methods perform poorly and contributed significantly to the evolution of the evaluation literature.

The methods employed to deal with selection in the evaluation literature can be classified into methods assuming that the selection occurs on observables and methods assuming it occurs on unobservables. The former methods, which comprise linear regression and matching methods, are used when participation is determined by characteristics that are fully observable to the evaluator (X). Therefore, by statistically conditioning on these

---

3 Also, random experiments are not immune to disadvantages thereby creating a need for the use of econometric tools. Disadvantages are discussed in Blundell & Costa Dias (2002) and Smith (2004), among others. Heckman & Lalonde & Smith (1999) for technical details. One of the common problems Smith (2004) highlights is the problem of dropout within the treatment group in experiments. Two ways are employed in the literature to deal with the dropout problem. The first way is to re-interpret the estimated treatment effect as the intent to treat or in other words the impact of the offer of the treatment rather than its receipt. Another way is to adopt an estimate adjustment introduced by Bloom (1984). This adjustment requires availability of follow-up data for dropouts, full-participants and control group members (and assumes random assignment). Bloom (1984) focused on “no-show” dropouts where a participant who was assigned for treatment did not show up and hence was not exposed to the program. He also discussed a “partial-show” case in the context of multiple-treatment design. Heckman, Smith and Taber (1998) focus on alternative treatment impact parameters in the case of partial treatment. They use an instrumental variable approach to identify the treatment impact. Heckman, Hohmann, Smith and Khoo (2000) analyze dropout cases where participants dropout to pursue better alternatives to the program (they also looked at the substitution problem where control group members seek other programs). They provide an estimate adjustment to deal with dropout and substitution.

4 For a flavor of this research debate, please refer to Human Resources Development Canada HRDC (1998).
characteristics, the selection issue is resolved. Formally, these methods depend on the Conditional Independence Assumption (CIA) which implies that assignment into treatment is random conditional on X.

Since these methods require controlling for X, they are naturally linked to linear regression methods. Unlike linear regression, matching methods do not impose linearity and, instead, employ non-parametric and semi-parametric methods. In order to estimate the counterfactual, outcomes of an individual are matched with the outcomes of an individual in the comparison group that is most similar (in terms of observable characteristics)\textsuperscript{5,6,7}.

Selection on unobservables takes place when unobserved variables affect both participation and outcomes. Three major methods are used in this case: instrumental variables, Heckman selection procedure, and longitudinal analysis. Following is brief discussion of each method:

1. Instrumental variables techniques have been around in standard econometrics for a long time. An instrument in the context of evaluation is a variable that affects participation but does not directly affect outcomes. Discussions of instrumental analysis in evaluation can be found in Heckman & Lalonde & Smith (1999), Angrist and Krueger (1999) and Moffitt (1991). The typical problem of finding a variable that would serve as an instrument applies. In the literature, natural experiments and regression discontinuity design (RDD) are a popular source of instruments (Cobb-Clark & Crossley 2003)\textsuperscript{8}. Another source of instruments is created by using a variant of random assignment such as “randomized encouragement” design. Under this design, units in a voluntary program are randomly given an extra incentive to participate such as additional information on the program or a monetary incentive. Receiving this incentive is then used as an instrument (see Smith and Sweetman 2002 for a discussion).

2. The standard 2-step Heckman selection correction models can also be used to deal with selection on unobservables. In fact, there is a close equivalence between selection correction procedures and instrumental variable analysis (Cobb-Clark & Crossley 2003).

3. Longitudinal data analysis uses pre-program and post-program data in order to overcome selection on unobservables. Difference-in-differences (fixed-effects) is one of the main tools in this area. The idea is to difference out the unobserved differences between the treated and the untreated by comparing the pre-post

\textsuperscript{5} For a clear exposition of matching methods, see Blundell & Costa Dias (2002).
\textsuperscript{6} For a discussion of matching methods’ merits (relative to regression) and implementation issues, see Cobb-Clark & Crossley (2003).
\textsuperscript{7} An extension to the matching literature is to allow for multiple treatments.
\textsuperscript{8} Treatment in RDDs is a deterministic function of a continuous variable where treatment eligibility requirements results in a sharp break in the value of this variable (e.g., a score on a test or a given age). The program impact estimation involves the comparison of observations just above and just below the cutoff using regression techniques. This approach is now trendy. Discussion of the technique can be found in Heckman & Lalonde & Smith (1999). Examples of applications include Lee (2008), Harvey et al. (2000), Lemieux and Milligan (2007) and Black (1999).
change of treated units with that of the untreated units. The underlying assumption is that the unobserved differences are time-invariant\(^9\). For an introduction to longitudinal analysis in evaluation research, see Moffitt (1991)\(^{10}\).

Combining longitudinal methods with other methods is also an approach that has been used in the literature. For example, control variables can be added to the difference-in-differences estimator. Furthermore, difference-in-differences can be combined with matching methods by differencing the data first and then applying the matching method (see Heckman & Ichimura & Todd 1997, Heckman & Ichimura & Smith & Todd 1998 for details and Eichler & Lechner 2002 for an application). This was sometimes found to outperform regular matching (e.g., Smith & Todd 2004).

An important recent development of the literature is to allow for heterogeneous impacts across treatment units unlike earlier when program impacts were assumed to be homogenous. This means that an evaluator is not only concerned with estimating mean program impacts. It also alters the interpretation of standard econometric estimates. See Heckman, Smith and Clements (1997) for a technical discussion.

Overall, it is important to note that econometric evaluation methods rely on different assumptions to overcome the selection problem. They would produce the same results only if there is no selection issue involved. Therefore, when different methods yield different estimates, this does not mean they are incorrect. Estimator appropriateness is context-specific and depends upon the problem at hand, the data design, and the institutional details surrounding the program (Heckman, Lalonde and Smith 1999). There is therefore no one absolutely correct estimator.

The literature on evaluation is continuously growing and econometric evaluation knowledge has increased dramatically. Researchers are now aware of the relative shortcomings and virtues of the different approaches and are no longer looking to solve selection problems with a magic bullet estimator as such an estimator does not exist (Smith 2004).

C. Proposed Methodology:

The program we propose to evaluate aims to transform the lives of out-of school adolescent girls in rural Upper Egypt by providing them with safe public spaces to meet, learn and play, giving them a second chance for education, and helping them acquire the skills necessary to become active members of their communities. The program is housed in village youth centers that, although ostensibly open to all youth, had until the beginning of the program been exclusively male spaces. Its curriculum emphasizes literacy and life skills with special attention to reproductive health, livelihoods information, civic engagement, and sports. The program also works to develop a cadre of more educated young women from the same

\(^9\) The random-growth model does not assume time-invariant differences. Instead, it assumes that the change takes place at a constant rate during the analysis period. Regner (2002) found that it outperformed the fixed effects model. However, it requires more than one pre- and post- program data points.

\(^{10}\) Lott and Mustard (1997), Black and Nagin (1998) and Ayres and Donohue (2002) go into details of potential problems with the longitudinal approach (Smith 2004).
communities to serve as promoters and role models for the targeted girls. The project also
works with other “gatekeepers” in these conservative rural communities, such as parents,
brothers, and community leaders, to enable the girls to participate and change social norms
concerning girls’ life opportunities.

The population to be studied thus consists of all out-of-school adolescent girls who are
12-15 years old at the beginning of the program including, participants and non-participants in
intervention villages and girls residing in control villages in which no intervention is taking place.
We will also be gathering information from the girls’ parents and adolescent boys living in
the same households. These target groups will be first interviewed in a baseline survey to be carried
out prior to the stat of the intervention and then again at the completion of the 20-month
program. We intend to go back and re-interview the girls included in the study five years after
the completion of the program if resources can be obtained to do that.

The proposed methodology involves randomizing the selection of intervention and
control villages among clusters of pre-selected villages that fulfill certain basic criteria. Each
cluster will initially consist of about twelve villages, of which five will be selected as intervention
villages and two as controls. We expect to have a total of six clusters. The intervention villages
will then be mapped and all their households listed to identify households containing out-school
adolescent girls in the 12-15 age range. The baseline questionnaire will then be administered to
a saturation sample of all the eligible girls in the village. Once the baseline survey is completed,
recruitment into the program begins. Two classes of 25 girls each will be recruited into the
program in each intervention village on a first-come first-serve basis. Since participation into
the program is voluntary and can therefore not be randomized, we will have to deal with potential
biases from selectivity into the program in the data analysis. Depending on the size of the
control villages, we will either map and list the entire village or randomly selected segments that
will yield the desired number of out-of-school girls for the control group. At the end of the
intervention, the endline survey will be administered to the same sample interviewed in the
baseline, irrespective of their participation in the program.

The outcomes that we wish to gauge the impact of the program on are as follows:

(i) functional literacy, cognitive skills and continuing schooling:
We will use a simple test to gauge the girls’ level of functional literacy and cognitive skills.
We will also inquire about any prior involvement in schooling, including grade attainment,
and repetition. We will monitor the level of success in passing government-sponsored
literacy tests and rates at which girls are mainstreamed back into preparatory (middle)
schools.

(ii) attitudes about marriage and childbearing:
We will inquire about attitudes as to ideal age at marriage and views about decision-making
regarding marriage partners and timing of marriage. We will also inquire about ideal family
size and fertility intentions.

(iii) knowledge about nutrition, hygiene, and reproductive health,

---

1 Some of these basic criteria are the presence of a suitable youth center and girls’ preparatory school
within walking distance.
A series of questions will involve basic knowledge about nutrition, hygienic practices, knowledge of basic bodily functions, reproductive behavior and family planning.

(iv) attitudes about harmful traditional practices, such as female genital mutilation/cutting: FGM/C is very prevalent in rural Upper Egypt, with prevalence reaching into the high ninety percent range. We will attempt to gauge the impact of the program on the girl’s own circumcision status, but we do not expect big changes to occur since circumcision usually occurs around the ages of eligibility for the program. However, we expect to see large effects with respect to intention to circumcise one’s daughter.

(v) social isolation, peer networks, and participation in group or community activities: Typically, social isolation increases as girls transition into adolescence and approach the age of marriage in Upper Egypt. We will investigate the extent to which the program has contributed to reducing this social isolation by widening the girl’s circle of friends and the level of support she is getting from peers, and the extent to which she participates in group or community activities.

(vi) gender norms, gender-based violence, mobility, and agency, including attitudes about participation in sports and recreation activities: We developed an index of gender norms based on attitudes about women’s work outside the home, traditional female roles, sex preference for schooling, decision making within the household, son preference, choice of marriage partners, and decision about spending. We have a series of questions about attitudes about domestic violence and the circumstances under which a girl “deserves” to be beaten.

Data about most of these variables has already been collected in the pilot stage. We have baseline and endline data on 587 girls from 4 intervention villages and 2 control villages. In addition, we have baseline data from an additional nine intervention villages and 4 control villages in which the project is currently being carried out. A review of the literature on the impact of those short-term outcomes on longer term outcomes as these girls transition to adulthood will be conducted.

As highlighted above, one of the main problems we anticipate to encounter in evaluating Ishraq is the non-random selection into participation and also the non-randomness of the drop-out process as girls drop out of the program midstream. First, since this is a voluntary program, we expect that the girls that self-select into the program are more likely to have parents who are more motivated and committed to education especially that of girls. Second, an additional element of non-random selection may be related to the project recruitment process. The pilot phase of Ishraq used several recruitment techniques including word of mouth, public announcement, and household visits by Ishraq promoters. The latter was the most effective in attracting participants. However, this technique may have resulted in bias because the promoters may have been more likely to approach people belonging to their social networks of friends and family\(^\text{12}\). Additionally, the promoters may have been more likely to approach households in the mother village (because of geographical proximity) and less likely to access the more-disadvantaged households in satellite villages. To overcome selection issues, we plan

\(^\text{12}\) Since the promoters are educated young women, households belonging to their social networks may value girls’ education more than the average household.
to undertake various strategies corresponding to the different stages of the project implementation (recruitment, pre-program activities including baseline data collection and data analysis.

Drop out from the program is also likely to be non-random. Girls whose parents want them to marry early; rely on the income from their daughters’ paid work, those who have a greater need for the girls’ participation in domestic chores, and those have less commitment to their daughters’ education might be more likely to drop out before the completion of the program.

The following sections delineate the different measures that will be undertaken to deal with the issue of non-random selection into participation and drop-out. These relate to the recruitment process of program participants; the randomization of the selection of intervention villages, the collection of data on additional variables and different verification techniques of data quality, and data analysis and estimation strategies. We discuss each of these measures separately.

C.1. Recruitment Process of Program Participants

Based on our experience from the pilot phase, we will adopt an improved Ishraq recruitment strategy that will help in reaching and mobilizing local communities. For example, we will be using visual advertisements in high-traffic areas of the villages. More training would be given to promoters to improve their skills in communicating information on Ishraq. More importantly, in the scale-up phase, we are forming Ishraq village committees that would include community leaders such as village mayors and religious leaders. Moreover, we are relying on a local NGO to implement the project at the community level with technical assistance and oversight from the Population Council and its partners.13

In the pilot phase, girls were accepted into the program on a first-come first-served basis until the classes were filled. Some people came to know about the program and expressed interest in joining the classes after the start of the program. That is why randomization was not possible among the pool of applicants that were accepted as applicants arrived at different points in time including after the start date of the project.

In the scale-up, we would like to advertise for an Ishraq workshop/regISTRATION event to be held on specific dates prior to the start of the program. The purpose of this event is to reach a wider audience and to give more information to the community on details of the project including its start date. If people are informed about the nature of the program and the expected level of commitment, this may reduce the likelihood of future dropout from the program. It would also be an opportunity to gather information regarding how the event attendees came to know about Ishraq. This would help us in identifying which recruitment strategies worked and whether the attendees are related to promoters. We will be maintaining records of the event attendees and of those who register. We are also considering a pre-

13 We will prepare an exhaustive list of NGOs in the governorate and then choose one after a screening process. On the one hand, a local NGO will be more aware of local institutional details. On the other, village residents will be more accepting of a local intervention. Additionally, hiring local NGOs will contribute to local capacity building.
program trial period where girls attend classes and decide whether they would like to join the program upon its start.

C.2. Randomization Strategies

Selection of intervention and control villages will be randomized. Based on census data, all villages with high rates of school drop out will be listed and control and intervention villages will be randomly selected from these lists.

However, while we are hoping the additional strategies we will adopt will be effective in propagating information prior to the program, we are not sure that the strategies adopted above would fully ensure that we have complete waiting lists prior to the program start. Hence, in practical terms, we are not sure if we can undertake randomized assignment from the pool of registered girls. Furthermore, randomization will pose a practical difficulty in terms of monitoring because local implementers are not familiar with such a procedure and hence may not preserve random assignment. Additionally, randomization may be resisted by implementers and village residents.

We also thought of alternative randomization procedures in intervention villages. However, we think, they would pose practical difficulties related to the nature of the villages in Upper Egypt. One procedure we considered was to randomly choose girls from baseline data. However, as per the standard randomization procedure above, this may be resisted by local residents. In addition, the voluntary nature of the program remains. Those who got selected will not necessarily eventually participate and hence we will be estimating the “intent to treat” rather than the impact of the treatment on the treated. Similarly, a “randomized encouragement” procedure where households are randomly selected to be given a monetary incentive may result in other village residents wanting to get similar incentives. Also, if extra information in the form of house visits is assigned to given households, we cannot ensure random assignment is preserved because in village settings, people commonly share information through word of mouth and might wonder why some people got the encouragement and others not.

C.3. Data Collection

Based on pilot data collection, we expect to get high response rates in the baseline survey. This is not uncommon in village settings or in Egypt in general. For example, the Egypt Labor Market Panel Survey 2006, in which two of the research team members have been involved, had a non-response rate of less than 2%. In the earlier Ishraq data collection experience, we were able to get a high follow-up rate (92% of those interviewed in the baseline survey were re-interviewed in the endline survey). Additionally, we do not expect to face a significant attrition problem resulting from control group members leaving their villages as people in village settings tend to be rather immobile except for small percentages of girls who marry and move to other villages. As mentioned above, we will be keeping record of people who express interest in the program but do not join because classes are filled, these will be considered as a “waiting list” group.

We will be collecting data from all households in intervention villages to get information on the universe of eligible girls. Data collection in the baseline stage will involve a mapping of
both control and intervention village. Mappers divide the village into sectors of equal number of households. In intervention villages, all eligible girls in all sectors will be interviewed. In control villages, a random sampling of sectors will be undertaken. All girls in the eligible age groups will be interviewed in these randomly selected sectors. The endline data collection process will use the same maps developed during the baseline data collection process to visit the same households.

To double-check that we are getting accurate figures, we would be comparing the population figures we get from the survey with the corresponding age-group population figures from the 2006 Census.

There are additional variables we plan to include, some of which we plan to use as instrumental variables as indicated in the next section. The set of additional variables include: parents education, mother age at marriage, whether parents were child-laborers, whether older sisters were circumcised, siblings configuration, distance (and time) to youth center (both based on self-reporting and village mapping information), distance (and time) to the center of the village. We also plan to collect information from community leaders on seasonal work. Finally, while the pilot phase questionnaire included questions on past attendance of literacy programs, we will expand on these questions.

C.4. Data Analysis and Estimation Strategy

Preliminary Steps

The first step we will take is to compare the baseline characteristics (i.e., prior to treatment) of control village and intervention village members to ensure randomization at the village level was successful. This could be done by looking at variables means and standard deviations and by testing if they are statistically different. Examples of variables to be compared include: parents education, wealth level (based on asset ownership), schooling status and number of completed years of schooling.

To get a feel of the extent of selection into treatment within intervention villages, we will compare the same set of variables across participants, waiting list members and non-participants. In addition, we will fit probit models for participation into Ishraq to see which variables are significant.

As a preliminary examination of the program impact, we will compare the level of changes in outcomes across the following groups (treatment, waiting list, untreated in intervention villages and control village members). We will also carry out an intent-to-treat analysis by comparing all girls in intervention villages (irrespective of participation in Isharq) to all girls in control villages.

Estimation Strategy

We will have data on both pre- and post- program characteristics in control villages and intervention villages (whether girls are participating or not). The quasi-experimental nature of our data will allow us to have the outcomes measured in the same way for the treated and

---

14 This is not an uncommon step (see Behrman & Todd 1999, and Rouse 2004).
untreated units. This is an important recommendation in many papers in the literature and is not always attainable (e.g., when the comparison group is constructed statistically using administrative surveys).

To deal with selection, we use several models: intent-to-treat, instrumental variable and difference-in-differences models. We will use specification tests to compare among the models. Additional specification/models are also discussed.

Villages are randomly assigned into control and intervention villages. Therefore, results will be unbiased if we use an intent-to-treat approach where the outcomes of control village members are compared to intervention village members whether they were treated or not (i.e., the treatment explanatory variable would equal zero if the girl resides in a control village and 1 if she resides in an intervention village). The advantage of this model is that there is no selection issue. However, the effect of the treatment will be diluted as the model will include untreated girls when estimating the average treatment effect. The following models deal with selection within intervention villages.

As explained in the literature section, to use an instrumental variable approach, we need a variable that affects participation but that do not directly affect outcomes. We will use the distance between the girl’s home and the youth center (where the project is housed) as an instrument as it proxies for the costs of participation. We will additionally use the distance to the village center as an explanatory variable in both the participation and outcome equations to capture access to other services and the effect of the centrality of the household location. Identification will thus depend on the variability in the distance between the youth center and the center of the village. We will collect information on distance (or time) based on respondents’ answers and on mapping conducted by the data collection team.

The third model we use exploits the fact that we have longitudinal data. The difference-in-differences estimator can be easily implemented by using a group dummy, period dummy (pre- or post- program) and an interaction of the group and period dummies\(^{15}\). We will combine the differences-in differences with other methods in additional specifications. A “regression-adjusted” specification will control for observable characteristics in addition to the group and period dummies and their interaction. We are also considering combining difference-in-differences with the propensity score method\(^{16}\). Heckman, Ichimura and Todd (1997) found this model to perform best overall compared to other models they employed.

We are also considering looking at an additional dimension of comparison whereby the program impact estimate is based on the difference between the outcomes of treated girls and that of waiting list girls. The waiting list group serves as a good comparison group because it should be equivalent to the treated group in terms of motivation and valuation of girls’ education (since they applied for the program).

Though specification tests have been recommended in the literature to sort among alternative estimators and to test for the sensitivity of the results, they have not been widely

---

\(^{15}\) We cannot use the random growth model to avoid the time-invariance assumption the difference-in-differences estimator imposes. This is because the random growth model requires more data points.

\(^{16}\) There are user-written STATA codes to perform propensity score estimation (for example, “pscore” by Becker and Ichino).
used. Heckman and Hotz (1989) is the key paper in this area. Among the tests they propose are pre-program tests. The rationale for these tests is that, prior to treatment, treatment dummies should not affect outcomes for the control or the to-be treated. In other words, the null hypothesis for these tests is that the treatment variable coefficient is equal to zero when using baseline data. The limitation of these tests is that they assume that if a model successfully adjusts for pre-program differences, treated and untreated units will also do so in a post-program period (Heckman & Ichimura & Todd 1997).

D. Human subject concerns

Any research project that proposes to collect information from minors must clearly address issues of how to obtain informed consent from the respondents themselves and from their parents or guardians. We have developed solid procedures to do so at the pilot stage. Moreover, the Population Council has an Institutional Review Board and a well-elaborated process to protect human subjects. The protocol for this study, including all survey instruments and informed consent procedures will have to be submitted for review and approval by the Council’s IRB.

E. Consultation and Dissemination Strategy:

The Population Council firmly recognizes that the ultimate aim of development research and projects is to provide and disseminate evidence-based results, leading to the formulation of informed policies aiming at reducing poverty. Through the numerous projects implemented in Egypt and other countries by the Council, we have developed strong partnerships with various stakeholders including government institutions, academics and civil society organizations. For this project, the Council will adopt a two-way approach where all stakeholders will be consulted throughout the research process and will, in turn, be thoroughly informed of the research results through a spectrum of appropriate dissemination strategies employed according to need. Research results will also be readily available through the Council’s website and publications distributed on a wide scale.

Throughout the project process, we will be working closely with our Ishraq partners namely Save the Children, CARITAS and CEDPA. These organizations have been intensely involved with the Ishraq program for its inception and are therefore familiar with all the circumstances and difficulties that have surrounded the program. Further, these organizations have a long history in working with girl issues and thus have much to offer in terms of input on the life situation of deprived Egyptian girls.

The Council plans on contacting key stakeholders directly, formulating a research advisory committee comprising members of governmental institutions, academics and civil society organizations involved in relevant areas of research. This committee’s input will be of key importance throughout the project as it will help identify specific research issues and considerations. Results of the project will be shared widely through a dissemination workshop where researchers and activists interested of our topic of interest will be invited.

Policy makers will be involved, in the form of the National Council of Youth (NCY) and the National Council for Childhood and Motherhood (NCCM), being the institutions most involved in youth and child issues. As part of our committee, their representatives will be consulted throughout the research process on key issues facing girls such as Female Genital Mutilation, deprivation and unattainable rights. We will therefore gain further insight into why
girl-friendly policies and mandates are difficult to implement and will tap into potential ways to tackle these obstacles. According to a verbal directive developed by Ministry of Youth, girls have equal rights to boys to use youth centers for sports and activities. However, this directive has not been translated into action with most centers being exclusively male spaces. One of our specific targets towards promoting policies supportive of the girl child will be to work towards the enactment of this verbal directive through promoting research results to government institutions.

Academics are essential stakeholders to link with for our research purposes. Aside from them providing key input for our research, they represent a vital link between us and a wider audience. Academics will disseminate our research findings to their colleagues and students to expand the outreach of researchers interested in and informed about issues facing Egyptian girls. The Council will also use MCAPI as part of our dedication to applying and exposing Egyptian researchers to innovative and efficient methods of research.

Civil society’s involvement in the project is necessary to provide our research with input grounded in the field. The implementation of our program in the field was shouldered by our partner NGOs. They are more informed about the situation in Egyptian communities than any other stakeholder and their contribution is vital. Our research will adopt a participatory approach whereby NGO activists, community leaders and members where our study takes place will be informed of the purpose, the methodology and the results of our project. We hope to better understand the dynamics taking place in Egyptian villages and how we can work with these dynamics. Our ultimate aim will to nurture a better relationship between NGOs and community members, whereby each group better understands the difficulties faced by the other and how they can better interact.

F. Study Team

Ragui Assaad is the Population Council’s regional director for West Asia and North Africa. He directs a diverse portfolio of research and programs in six countries. The Council’s activities in the area are aimed at building leadership and research capacity, improving the quality of care for reproductive health and family planning services, reducing poverty, enhancing civil participation, and supporting transitions to adulthood. Assaad’s recent research focuses on labor market analysis, gender inequity in the labor market, and work and family issues in the Middle East and North Africa. His studies examine gender inequities in economic restructuring, poverty and poverty alleviation strategies, the informal economy, community development, and urban planning in the developing world.

Over the past 14 years, Assaad has collaborated with Council staff in Cairo and New York on a number of projects relating to youth livelihoods, child labor, woman’s paid work, and the empowerment of adolescent girls. He has served since 1990 as a professor at the Humphrey Institute of Public Affairs, University of Minnesota. Since 1994, Assaad has been a research fellow of the Economic Research Forum for the Arab countries, Iran, and Turkey, an independent regional network of economists, policymakers, and business professionals. Assaad has published over 30 articles in peer-reviewed journals and edited volumes. He edited a book on the Egyptian labor market and contributed to a number of official reports published by
international organizations and think tanks. He has given numerous presentations at conferences and workshops around the world.

The holder of a B.Sc. in physics and an M.Sc. in mechanical engineering from Stanford University, Assaad also earned a Ph.D. in city and regional planning from Cornell University.

His role in the project will be to provide intellectual leadership to the team and to participate in data analysis.

**Ghada Barsoum** is the Senior Program Manager for the Poverty, Gender and Youth Program at the Population Council Cairo office. Barsoum has an extensive research experience in the areas of youth employment, poverty, micro-finance and education. She has been Project Coordinator of the Egypt Labor Market Panel Survey (ELMPS06) and was responsible for the oversight of data collection, data processing and documentation. She has published a book on the employment crisis of female graduates in Egypt by the American University in Cairo and four papers in peer-reviewed journals. She received her PhD from University of Toronto, Department of Sociology in 2005 and her MA from the American University in Cairo in 1999.

Her role in the project will involve the overall project management; participation in the design of questionnaire; oversight of data collection; analysis and dissemination activities.

**Rania Roushdy** is a Program Associate at the Population Council’s Cairo office. She is an Economist with main interest in Applied Econometrics and Economic Development. Roushdy joined the Population Council in June 2003, directly after she got her Ph.D. from the State University of New York at Stony Brook. Her current research lies in four areas: the linkage between urban poverty and health seeking behavior with a focus on social learning; poverty mapping and small area estimation techniques; intra-household allocation of resources, women empowerment and investment in children; and fertility behavior and childbearing attitudes. Roushdy has participated at numerous international conferences and has several working papers on: urban poverty and neighborhood effects on schooling attainment, and child health and mortality; effect of women status within the household on child investment; fertility preferences and demand for children; and measurements of human security through a knowledge based decision making system. Some of these papers are currently in preparation to be submitted to international journals.

Her role in the project will involve the participation in the design of questionnaire; data analysis and dissemination activities.

**Asmaa El Badawi** is currently a research fellow at the Population Council. She expects to defend her dissertation in Summer 2007 at McMaster University, Department of Economics. She holds a Master’s degree in Economics from University of Waterloo in Canada. She has done extensive
research in the area of education in Egypt. Her role in the project will involve data analysis and participation in questionnaire design.

Nadia Zibani is a research coordinator at the Poverty, Gender and Youth program at the Population Council. She is currently involved in projects related to youth in Egypt (out-of-school adolescent girls and child labor in the SME sector in Egypt).

Prior to the Population Council, Nadia worked extensively on child labor issues. Involved in the Egyptian Labor Market Survey of 1998, she conducted research on children’s work and schooling in Egypt in collaboration with Dr. Ragui Assaad and Dr. Deborah Levison at the Humphrey Institute of Public Affairs, University of Minnesota. In the early 90s, she conducted research in the areas of labor, employment and poverty for the Social Fund for Development and the Ford Foundation and participated in the design and development of a survey to determine the linkages between poverty and employment in the Metropolitan Urban Cairo Region.

At the Population Council, Nadia has been heavily involved in the Ishraq program since in both research and programmatic sides including conducting qualitative research with secondary target groups/gatekeepers (parents, and brothers) of the beneficiaries in the four intervention villages; coordinating the sports component; designing a sports curriculum tailored for rural adolescent girls; working in close collaboration with the Egyptian Ministry of Youth (MOY) and other sports federations (national and international) in developing guidelines for the implementation of the scale up of the Ishraq sports component; and the overall design, implementation and coordination of the monitoring and evaluation activities.

Nadia holds Diplôme d’Etudes Approfondies (DEA) in Economics Development from the University of Paris I, Panthéon-Sorbonne.

Her role in the project will involve the participation in the design of questionnaire; oversight of data collection; analysis and dissemination activities.

G. Timeline

The timeline to this study will slightly deviate from the one included in the guidelines. This is mandated by the fact that the intervention program that is being assessed lasts for twenty months and is due to start in the field in August 2008. Following is the timeline of the study based on this early start of the project:

2. Dates for the submission of interim reports, presentations of survey results will follow the timeline listed in the guidelines
3. Program intervention: August 2008 – April 2010
4. End-line data collection: February 2010. This date is mandated by PEP’s timeline for deliverables. While program activities will not be finalized by then, girls will have been exposed to a great proportion of the program.
5. Draft final report: June, 15, 2010
6. Presentation of draft report at PEP meeting: June 2010
7. Final report: July 31, 2010
8. Working papers: August-December 2010
Annex to Proposal in Reply to Reviewers’ Comments

May 23, 2008

This annex responds to comments raised by PEP as posted in April 2008.

1. **Re discussions with Relevant Policy Makers**

   In its current scale-up phase, Ishraq will be implemented in full partnership with a governmental partner, the National Council of Youth (NCY). The resource team of four international non-governmental organizations who implemented the pilot phase is now confident that the experience is ready for wider replication. This will be accomplished through the transfer of the model ownership to NCY to be the primary implementer of the project in order to ensure its sustainability. Data collected as part of the earlier pilot phase and the ensuing analysis have been instrumental in ensuring the buy-in of this national partner of the model. Analysis of this earlier data set has been widely disseminated and was also published in order to make the case for the value of programming addressing out-of-school girls and excluded adolescents in rural Upper Egypt, a geographical area with a concentration of poverty.\(^\text{17}\) NCY is convinced of the value of research and the documentation of program impact as it transitions into its new phase.

2. **Re Team Composition and Capacity Building**

   The Council’s team is ideal according to the criteria for team composition as shown in the guidelines. Guided by an experienced researcher, Dr. Ragui Assaad, the team is comprised of young researchers who have either recently finished their PhDs or are in the process of obtaining their degrees. All researchers, except for the team leader, happen to be females. One of the two researchers attending the training provided by PEP, Asmaa Elbadawy, is less than thirty years old. She is a trained economist in the final stages of the PhD program at McMaster University. The project will add to her expertise on girls’ education, an area she has been working on for her dissertation. Moreover, impact assessment is new to Asmaa and hence she would significantly benefit from the training. The other participant, Ghada Barsoum, obtained her PhD in 2005 from University of Toronto. Though a trained sociologist, she has little experience with quasi-experimental research methods. She manages the Poverty, Gender and Youth Program and the Council’s WANA office, which implements the Ishraq program.

---

intervention and carries out its impact assessment. Training will be of great benefit to her and to our organizational capacity.

Research activities supported by this project will be the first program evaluation attempt using advanced econometric techniques at the Population Council’s WANA office in Cairo. While efforts were made to evaluate the pilot phase of Ishraq, the sample size did not allow for undertaking advanced analysis. As such, there is a capacity building element for team members involved in the project. Various aspects in the scale-up phase are new to our team. First, the measures that we plan to introduce to minimize self-selection into the program (e.g., Ishraq village committees and the Ishraq workshop/registration event) will entail some capacity building for some team members as well as the local NGOs and Ishraq promoters.

Using econometric impact evaluation methods will involve capacity building in terms of acquiring a good knowledge of the evaluation literature and an understanding of the techniques that will be employed to deal with self-selection and non-random dropout as well as their implementation using a statistical package (in our case, Stata).

3. Re Human subjects concerns

As noted in the proposal, the Council has an Institutional Review Board (IRB) to ensure the adherence to ethical standards in the collection of data from minors. The IRB is an oversight committee, whose mandate is to protect the rights and safety of the subjects of the Council’s research. Ghada Barsoum is currently in the process of getting an Ethical Training Certificate covering issues of confidentiality, risk and benefits to interviewees and informed consent. These standards are strictly followed by research done by the Council.
References:


