Janssens, Wendy

Working Paper

Measuring Externalities in Program Evaluation

Tinbergen Institute Discussion Paper, No. 05-017/2

Provided in Cooperation with:
Tinbergen Institute, Amsterdam and Rotterdam

Suggested Citation: Janssens, Wendy (2005) : Measuring Externalities in Program Evaluation, Tinbergen Institute Discussion Paper, No. 05-017/2, Tinbergen Institute, Amsterdam and Rotterdam

This Version is available at:
https://hdl.handle.net/10419/86210

TI 2005-017/2
Tinbergen Institute Discussion Paper

Measuring Externalities in Program Evaluation

Wendy Janssens

General & Development Economics, Faculty of Economics and Business Administration, Vrije Universiteit Amsterdam, and Tinbergen Institute.

Tinbergen Institute
The Tinbergen Institute is the institute for economic research of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam, and Vrije Universiteit Amsterdam.

Tinbergen Institute Amsterdam
Roetersstraat 31
1018 WB Amsterdam
The Netherlands
Tel.: +31(0)20 551 3500
Fax: +31(0)20 551 3555

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Please send questions and/or remarks of non-scientific nature to driessen@tinbergen.nl.
Most TI discussion papers can be downloaded at http://www.tinbergen.nl.

Measuring Externalities in Program Evaluation

Spillover effects of a women's empowerment programme in rural India

Wendy Janssens[1]

March 2006

Abstract

Impact evaluations of development programmes usually focus on a comparison of participants with a control group. However, if the programme generates externalities for non-participants such an approach will capture only part of the programme's impact. Based on a unique large-scale quantitative survey we estimate the direct as well as the spillover effects of a women's empowerment programme in rural India on child immunization and school enrolment. The survey covers both participants and non-participants living in programme villages, as well as respondents in control villages where the programme is not yet active. We account for participation selection bias using instrumental variables. The control villages allow us to test the exclusion restriction and provide us with an effective control group to analyze programme impact. We find both direct effects and significant spillovers on non-participants. The impact of interventions might be substantially underestimated if such external effects were not taken into account.

Acknowledgements. The author is very grateful to Jacques van der Gaag and Jan Willem Gunning for their support and many detailed comments. She greatly appreciates the discussions with Robert Sparrow and Chris Elbers. Thanks also go to Jaap Abbring, Richard Blundell, Martin Browning and other seminar participants at the Tinbergen Institute, the Network of Excellence COST Conference, CERES, the Netherlands Ministry of Foreign Affairs and the World Bank. Special thanks are due to the Mahila Samakhya programme in Bihar for its cooperation and hospitality, and to Shantwana Bharti for the many inspiring discussions. This research was financially supported by the Early Child Development (ECD) team of the World Bank.

[1] Wendy Janssens is with the Vrije Universiteit Amsterdam, the Amsterdam Institute for International Development and the Tinbergen Institute.

1. Introduction

Impact evaluations are generally based on a comparison of the treatment group with a control group. However, if the project generates externalities and the control group benefits from the programme as well, such an approach might seriously underestimate the programme's impact. We examine the impact of Mahila Samakhya, a women's empowerment programme in the North-Indian state Bihar, on child outcomes. We focus explicitly on the spillover effects of the programme on children of non-participants who live in a village where the programme is active.

Bihar is one of the poorest states of India with very low educational and child health indicators. It is a predominantly rural state with a large and rapidly growing population of almost 83 million people. Society is very caste-conscious and gender biased, and still marked by the former feudal practices of bonded labor. The poor social conditions in Bihar are accompanied by severe material poverty for a majority of the population. Almost two thirds of all people, 64%, lived below the poverty line of US$1 per day in 1993/1994, a very large proportion compared to the national average of 36% (World Bank, 1997). In addition, the population is confronted with a poor state of infrastructure and social services that exacerbate the difficult circumstances.

The strong social stratification and the gender bias result in substantial disadvantage in primary school access and learning for children from the lowest castes and for girls. Enrolment is low, drop out and retention rates are high and learning achievement of those completing primary school is very low (World Bank, 1997). This leads to the lowest literacy levels of the entire country. The overall literacy rate in Bihar is only 48% compared to the national literacy rate of 65%[2]. Disaggregation by gender shows that 60% of the male population in Bihar can read and write, but only 34% of the female population. The immunization rates show a similar picture. In 1998/1999, only 11% of all Bihari children aged 12 to 23 months were fully immunized compared to a national full immunization rate of 42%[3].

Under these circumstances, the Mahila Samakhya programme (literally `women interacting as equals') aims to educate and empower women in rural areas, especially women from socially and economically marginalized groups, such as women belonging to the Scheduled Castes[4]. The women are encouraged to set up women's groups in their communities to improve their own lives. The programme offers training and support to the groups. Improving the education and health of children are important components of the group activities.

This paper evaluates the impact of the Mahila Samakhya programme on child outcomes in three districts of Bihar.[5] The evaluation is based on a unique data set

[2] India Census Office (2001a, 2001b). Literacy rates are calculated over the population aged 6 years and older.
[3] National Human Development Report 2001 (Government of India, 2002).
[4] Scheduled Castes is a classification in the Indian Constitution. 'Dalit' or 'Harijan' are other often used terms to refer to the Scheduled Castes. Many of these people were considered to be 'untouchables' in India for centuries (Ramachandran, 1998, p.71).
[5] For a recent overview of impact evaluations of community-based development projects, see Mansuri and Rao (2004).

encompassing 1991 women in 102 villages. In 74 of these villages the programme is active, in the remaining 28 villages the programme has not started yet. These latter villages will be used as control group. Due to the slow scaling up of the Mahila Samakhya programme, the matched control villages are highly similar to the programme villages, thereby providing for a good counterfactual.[6]

Within the programme villages, both participants and non-participating women were interviewed. This sample design allows us not only to compare the outcomes of participants with those of non-participants in the programme villages, but also to examine the externalities of the programme through a comparison of non-participants in programme villages with the control group. Although a comparison of child outcomes at the village level captures the overall programme impact within the community, it will not provide insights in the potential presence and size of externalities.

A growing body of literature emphasizes the role of social interactions in shaping environments and outcomes. There is increasing evidence of the influence of neighborhoods and peers on outcomes such as child health and behavior (Case and Katz, 1991; Katz et al., 2001), student outcomes (Lalive and Cattaneo, 2004; Sacerdote, 2001), technology adoption in agriculture (Besley and Case, 1994; Foster and Rosenzweig, 1995; Munshi, 2004), retirement plan decisions (Duflo and Saez, 2002) and contraceptive prevalence (Munshi and Myaux, 2002).

However, very few impact evaluations of development programmes explicitly take into account the potential presence of local spillover effects.[7] For example, Alderman et al. (2003) and Kim et al. (1999) measure the increased enrolment of boys after the introduction of girls' primary school programs in Pakistan. Bobonis and Finan (2002) measure increased enrolment of children from households not eligible in the PROGRESA programme in Mexico. Miguel and Kremer (2004) measure externalities of a deworming programme in Kenya due to reduced disease transmission on children attending neighboring schools. These evaluations use the random characteristics or the eligibility criteria of the programme to identify the spillover effects.

However, when implementation is not random or participation is voluntary and open to all, identification of the treatment effect becomes more problematic (Manski, 1993; 2000). It is usually difficult to plausibly argue that the exclusion restriction for the instrumental variables holds. In this paper, we propose a different test of the exclusion restriction that exploits our survey design. The control group allows us to test directly whether the instruments can be excluded from the outcome equation. An estimation of the treatment effects using propensity score matching provides additional support to the results.

The results show that the programme has a significant positive impact on immunization rates and preschool and school enrolment rates when one of the female household members participates in a Mahila Samakhya women's group. Moreover, the spillover effects on children living in programme villages, but whose mother does

[6] Duflo (2003) discusses the potential contribution of scaling up to programme evaluation design.
[7] Impact evaluations that account for general equilibrium effects are more common. For an overview of analyses of general equilibrium effects in training and labour market policies, see Heckman et al. (1999).

not participate herself, are substantial, especially for girls and children belong to the Scheduled Castes. These results emphasize the importance of using the right control group in an impact evaluation, especially if the programme focuses on awareness raising and collective action. A comparison of participants with the control group only, would have missed to measure the spillover effects within the village. The omission of the control villages in the survey would have led to a double underestimation of the programme effects because the outcomes of participating households would have been compared with the improved outcomes of non-participants.

The structure of the paper is as follows. The next section describes the Mahila Samakhya programme and its objectives. A comparison of participants with non-participants in programme villages shows that the programme reaches especially those women that belong to marginal groups in society. The third section discusses the sample selection and data collection procedures. It also compares the characteristics of programme villages and control villages in the sample and argues that the control villages form a good counterfactual. The fourth section gives a description of the primary statistics and differences in immunization and enrolment rates for the different sample groups. Fifth, we will outline the estimation model and examine the underlying assumptions implicit in the various estimators of the programme effects. In the subsequent section we will discuss the econometric results of both the two-stage instrumental variables estimation and the propensity score matching. The final section concludes.

2. Description of the programme

2.1 Programme objectives and characteristics

The Government of India launched the Mahila Samakhya programme in 1988 to support its 1986 New Policy of Education. The programme is currently active in nine states[8]. The Mahila Samakhya programme was introduced in the state of Bihar in 1992. In March 2002 the programme was operational in 7 out of the 37 districts of Bihar. 1890 women's groups were active, covering more than fifty thousand women.

Because of its inclusion as a component in educational programmes, Mahila Samakhya in Bihar initially aimed at educational issues. The participating women obtained literacy and numerical skills, were motivated to increase school enrolment of their children, especially their daughters, and to become active members of Village Education Committees.

The Mahila Samakhya programme emphasizes the process of empowerment, instead of the fulfillment of targets. It seeks to bring about a change in women's perceptions about themselves and their own abilities, and the perceptions of society with respect

[8] The national Department of Education organizes the Mahila Samakhya programme at the national level. At the state levels, the programme is implemented through autonomous Mahila Samakhya Societies. Mahila Samakhya in Bihar started as a component of the Unicef-sponsored Bihar Education Project (BEP). From 1998 to 2003, the programme in Bihar was financed through the World Bank-sponsored Third District Primary Eduction Project (DPEP III) and currently it is part of the national Sarva Shiksha Abhijan programme (Movement for Education for All). The national Department of Education organizes the Mahila Samakhya programme at the national level. At the state levels, the programme is implemented through autonomous Mahila Samakhya Societies.

to women's traditional roles.[9] The programme does not prescribe the kind of activities that a group has to engage in. Instead, it aims to assist the women in their own solutions for problems. It offers training in a large number of areas and enables the groups to effectively access and utilize resources and government subsidies. As a result, the groups have taken additional initiatives over time to address a wide range of issues. These range from meeting daily minimum needs through savings and credit groups or income-generating activities, obtaining better health and hygiene knowledge, entering local politics, improving village level infrastructure, settling conflicts in the village, and setting up informal preschools and primary schools for girls (Mahila Samakhya, 2002).

The facilitators of Mahila Samakhya are responsible for the initial mobilization of women into a group. This is a slow process. The Mahila Samakhya programme does not offer clear-cut and immediate benefits. These should be produced by the groups themselves, albeit with support of the programme. Moreover, for many villagers the idea of women's empowerment seems in conflict with their traditional way of life. Through regular visits to the women's homes, the fields and their work place, the facilitator tries to gain trust among the women. During these visits she talks and listens to them and assists in solving small issues. The initial discussions in the group are for many participants the first time that they reflect on their situation and articulate their needs (Mahila Samakhya, 1995). As the group gets stronger, it starts to set its own agenda and meeting times, and the facilitator will participate less in meetings. This process usually takes 6 to12 months.[10]

2.2 Comparison of participants with non-participants

The group sizes in the three districts under study range from ten to eighty members with an average of twenty members. This represents between 1% and 26% of the total adult female population in a village (5.4% on average). In principle every woman who is interested can join the women's group. However, the facilitators gear their efforts particularly towards poor women from the Scheduled or Other Backward Castes. Table 1 provides a description of the population and community variables.[11] As table 2 indicates the majority of the participants belongs to the target group. Of the participants, 39% belong to the Scheduled Castes and 49% to the Other Backward Castes. Only 2% of the participants come from General Castes. The percentage Muslim population is a bit higher among non-participants with 11.6% as compared to

[9] A large number of women's empowerment programmes aim to address the disadvantaged position of women in Indian society. See for example Patel (1998) for an analysis of the relationship between the women's movement and women's education. Lahiri-Dutt and Samanta (2002) provide a discussion of the effectiveness of state initiatives on rural women's empowerment. Purushothaman (1998) gives a thorough case study of an informal network of NGOS and women's collectives, comparable to the Mahila Samakhya women's groups, in the Indian state Maharashtra. Agarwal (2001) analyses how seemingly participatory institutions and community groups in India can nonetheless exclude women to a significant extent.

[10] Data are calculated based on the data collection at the Mahila Samakhya district offices in 2003.

[11] All descriptive statistics take into account the survey design through appropriate weighting of observations. We correct for the stratification in weights and standard errors. Moreover, weights and standard errors are adjusted for the clustering at the village level and the fact that villages were not sampled proportionate to size. Finally, appropriate weighting of the households takes into account the disproportionate sampling of participating and non-participating households in programme villages.

9.2% of the participants although the difference is statistically insignificant. The rest of the population is in large majority Hindu.

The average participant has lower income than the average non-participant, and lower average land ownership. Participants live in houses with walls, floors, and roofs of lesser quality. They have less access to private water facilities or electricity. They more often use biomass as fuel for cooking. Participants also have less access to adequate sanitation facilities. However, the availability of sanitation is alarmingly low overall. Only 17% of the population in the study area have access to sanitation facilities such as a pit toilet or a latrine.

As table 2 shows, participants live in households with on average significantly less education than non-participants. Participants have slightly fewer household members but more children under the age of 14, resulting in a larger child dependency ratio for participants than for non-participants. The percentage of female-headed households is equal among the two groups.

In sum, women who decide to join a women's group belong to the poorer and socially more disadvantaged segments of society. In terms of child outcomes, we would expect their families to have lower enrolment and immunization rates relative to non-participants. On the other hand, participation in the programme is voluntary. There might be unobserved differences between participants and non-participants in terms of motivation, awareness or ability that both affect the likelihood that a woman participates and the education or health outcomes of her children. A comparison of participating households with non-participating households, even when controlling for observed characteristics such as income or caste could potentially be biased due to this self-selection into the programme.

Participation selection bias is a key problem in impact evaluation, and will be corrected for using instrumental variables. However, before the programme starts the awareness about child development is very limited overall and especially among the target group.[12] Given the fact that it takes about nine months for the facilitators to convince women to organize, it seems likely that their interest in child issues increases mainly *after* the first contact with Mahila Samakhya officers. Nonetheless, we will correct for potential individual differences in motivation.

3. Sample selection methodology

3.1 Sample selection

The study encompasses three districts in the north of Bihar: Sitamarhi, Muzaffarpur and Darbhanga. In these districts, the Mahila Samakhya programme has been active since 1992, 1993 and 1998 respectively. In each of the districts, groups exist that have just started and groups from the early years.

Of the three districts, Muzaffarpur is closest to the capital and relatively well connected by rail and road. It is one of the centers of political activity in Bihar. It has the highest overall and female literacy rates of the three districts (48.2% and 35.2%

[12] Based on interviews with Mahila Samakhya programme officials in October 2002 and during the field survey from March to June 2003.

respectively). It also has the highest percentage of Scheduled Castes population. Sitamarhi, south of the border with Nepal, is the most rural of the three with highest population growth, lowest sex ratio and lowest literacy rates (39.4% overall and 26.4% female). The district is relatively isolated and suffers from severe floods every year, resulting in seasonal mass migration. Darbhanga's situation lies somewhere in between. It is better connected than Sitamarhi, suffers more from floods than Muzaffarpur and has an overall literacy rate of 44.3% and a female literacy rate of 30.4%. Darbhanga has a relatively large Muslim population compared to the other two districts (India Census Office, 2001b).

The survey is based on a two-stage stratified clustered sample design. The sample consists of six strata: programme and control villages in each of the three districts. In the first stage, we randomly selected 25 programme villages in each district from the list of programme villages and 10 control villages in each district from control blocks[13]. These villages are the 'clusters'. In the second stage, we randomly selected 20 households in each village. To ensure a sufficiently large sample of participants in each village, we randomly selected 10 non-participating households in programme villages and 10 participating women. In control villages, 20 households were selected. The final sample consists of 74 programme villages and 28 control villages, 718 respondents participating in the programme, 714 respondents living in programme villages but not participating themselves (nor anyone from their household), and 559 control respondents living in villages where the programme is not active.[14]

The Mahila Samakhya programme usually starts in a few blocks per district. Only after these blocks are (almost) fully covered, that is, when facilitators have visited almost all villages, the programme expands to new blocks in the district. The programme focuses on districts and on blocks within districts that are particularly disadvantaged. The general selection criteria of Mahila Samakhya for programme blocks are threefold: a low level of female literacy, a high percentage of population living below the poverty line, and a high percentage of population belonging to the Scheduled Castes. Eventually, the programme aims to cover all blocks entirely. Table 3 gives an overview of the number of blocks and villages covered in the three districts.

The Muzaffarpur district office follows the general selection rule for blocks. In Sitamarhi on the other hand, the programme has started in most blocks at the same time, but it is active in only a limited number of villages in each block. Darbhanga is another exception. The programme started only recently in this district. Based on travel considerations it chose the two blocks that are closest to the district capital.

All program blocks in Darbhanga and Sitamarhi are selected in the survey sample. In Muzaffarpur only the three initial blocks are included, as the other two blocks were entered very recently. As control blocks, we choose blocks that are comparable in

[13] The administrative structure divides States in districts, districts in blocks, and blocks in revenue villages. Revenue villages cover one or a few villages dependent on their size.

[14] Three villages and an additional 49 interviews were dropped from the sample because of incomplete or unreliable data. This represents 3% of the sampled villages and 5% of the sampled households. In these villages, the supervisor of the interviewers had not been present during the survey resulting in problems with the questionnaires. It is not likely that the omission of these particular observations from the sample leads to a sample bias.

selection criteria to the programme blocks in each district. In Muzaffarpur, the two control blocks are relatively similar in the key selection indicators to the programme blocks. In Darbhanga, we choose the two next closest blocks to the district capital to serve as control group. For Sitamarhi, the three blocks where no or only a few groups have been started so far will serve as a control.

Within programme blocks, all villages are eligible for the programme. The process of village selection is as follows. The Mahila Samakhya programme announces the vacancy of facilitator jobs. Facilitators start with the mobilization process in their own village and four neighboring villages. After one year, they expand their working area to five additional villages in the block that score high on the Mahila Samakhya selection criteria. Due to difficult travel circumstances, in Darbhanga all ten villages lie close to the facilitator's own village.

3.2 Data collection

Data collection occurred in two phases. A first exploratory field visit took place in October 2002. During this visit Mahila Samakhya officials and facilitators were interviewed as well as women's groups in the field. These interviews were qualitative and served as inputs into the design of two questionnaires: one household questionnaire and one village questionnaire. The village questionnaire contains a general part on village facilities and characteristics, and a separate module on the women's groups. These were administered to the women's groups. In control villages, village leaders and other community members were interviewed for the first general village section only. After a pilot of the questionnaires, the actual data collection took place from March to June 2003.

3.3 Comparison of programme and control villages

To what extent do the control villages represent a good counterfactual for the programme villages? As described in section 3.1 the control villages are chosen from blocks that are comparable in selection criteria to the programme blocks. The slow scaling up of the programme within districts indeed offers the opportunity of constructing an adequate control group, as we will argue in the current section.

A comparison of the village characteristics in these control and programme blocks is given in table 4. Panel A of the table looks at population characteristics within control and programme villages respectively. These relate to household caste and religion, household income and wealth measures, household education variables and characteristics regarding the composition of the household. The differences between the two groups are small and highly insignificant for all characteristics.[15] Note that the recruitment method for facilitators could potentially have favored the villages with a higher female literacy rate in the programme blocks, since those villages might be more likely to provide suitable applicants for the job. However, we do not find a systematic difference in female education between programme and control villages.

[15] The results in table 4 panel A are based on a comparison of the total population in programme villages with the total population in control villages. Results are highly similar when we first calculate the average household characteristics per village and then compare the programme village averages with the control village averages.

Both show an average female education level in the household of less than primary school incomplete. Also the pre-programme 1991 village level female literacy rates are virtually identical among the programme and control villages with rates of 14.4% and 14.2% respectively (India Census Office, 2001a).

An examination of village characteristics in Panel B shows a similar picture. Again the differences between programme and control villages are small and insignificant. Control villages are slightly more likely to have experienced a flood in the last three years. Differences regarding the occurrence of draughts, the quality of the roads, and availability of public transport are insignificant.

The only notable exception is distance to the nearest town, which is considerably larger for control villages. The significantly larger average distance for control villages is probably a result of the programme block selection criteria. Both travel distance and distance to the nearest town increase if the villages in a block are more dispersed and remote. However, the larger distance to a town does not seem to affect the overall access to facilities that are generally considered to be relevant indicators of village level development. The percentage of villages with a market, post office, telephone, bus stop, bank or health center within their boundaries is about equal for programme and control villages. Likewise, in those cases where the facility is not available, we do not find any significant differences in the distance to the nearest facility outside the community.

Finally, panel C of table 4 compares the presence of schools in programme versus control villages. With the exception of preschools, the percentage of villages with at least one primary, middle or secondary school available is roughly comparable between the two groups. As expected, preschools are more numerous in programme villages since the Mahila Samakhya groups have set up a considerable number of preschools. Note that the women's groups have started an equally large number of informal primary schools for girls. However, these do not make a difference in the primary school statistics since in most of those villages a formal primary school exists as well.

Overall, the descriptive statistics show that the programme and control villages are highly comparable on a large number of population and village development characteristics. Thus, it seems plausible to assume that in the absence of the programme, child outcomes in the programme villages would have been similar to those in the control villages. The rest of the paper will assume that the control group represents a good counterfactual for the programme villages.

4. Descriptive statistics

In measuring the impact of Mahila Samakhya we focus on the effects of the programme on children of participants and non-participants in programme villages. In particular we will examine the effect of mother’s participation on child immunization and preschool/school enrolment, as well as the externalities on children of non-participants. Girls and Scheduled Castes children will receive special attention in the remainder of the paper. First of all, they are the main target groups of the programme.

Second, these children are especially disadvantaged on a large number of issues, including education and health (Drèze and Sen, 2002).[16]

4.1 Immunization

We study the impact of the programme's activities on four vaccines against six types of preventable diseases: BCG (Bacille Calmette-Guérin vaccine against tuberculosis), polio, DTP (against diphtheria, tetanus and pertussis or whooping cough) and measles.[17] All four vaccinations are free, but they are distinct from each other in several respects. To eradicate polio, the national government has launched a large-scale immunization campaign in 1995, the Pulse Polio Immunization Programme, to immunize all children under five against polio. At a few years intervals thousands of health workers are sent to all parts of the country to administer polio drops. Local community groups such as Mahila Samakhya groups are encouraged to participate in the organization of the campaign. Tuberculosis is the only vaccination that is not administered locally but at the block hospital. DTP and measles vaccinations have in common that they are both due in a relatively short period after birth. Both the DTP and the polio vaccine consist of several doses. Finally, immunization against measles suffers from the additional problem that it conflicts with a traditional belief. A goddess is believed to inflict measles on the family and child. Circumvention through immunization might anger her and cause even larger health problems. Especially in very traditional remote areas, this could hamper the increase of the measles immunization coverage rate.

The programme offers health education through a variety of means. First of all, the programme directly disseminates information to the women's groups through the weekly or monthly visits of the facilitator. The programme also organizes health trainings at the district level for the women's groups. To improve the quality of these health camps, Mahila Samakhya cooperates with health department functionaries such as Auxiliary Nurses and Midwives (ANMs), Lady Health Visitors and government doctors. In addition, the Mahila Samakhya programme in Sitamarhi and Muzaffarpur is engaged in a partnership with the World Health Organization (WHO) to train barefoot health workers. A representative of the women's group participates in the WHO training and disseminates the newly gained information about health and nutrition on her return in the village. The trainings increase health knowledge and

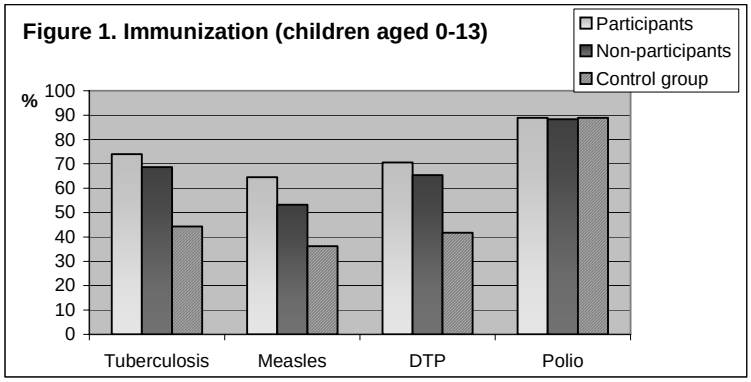
[16] See Rao et al. (2003) for a socio-cultural analysis of gender and caste differences in primary education in India, Thorat (2002) for a discussion of current day discrimination of Scheduled Castes, and Hoff and Pandey (2004) for an experimental investigation of deeply ingrained beliefs among Scheduled Castes elementary school children. Kishor (1993) and Das Gupta (1987) study gender differences in child mortality and the underlying mechanisms that bring about these differentials. Pande (2003) analyses the gender inequalities in childhood nutrition and immunization in rural India. Whereas the differences in immunization between boys and girls are diminishing over time, the disadvantage in child health, child mortality and immunization coverage of Scheduled Castes children compared to other castes remains strongly present (International Institute for Population Sciences, 2000). The World Bank (2001) and Filmer et al. (1998) provide overviews of gender inequality research, also covering India.

[17] In 1978 the Government of India initiated the Expanded Programme on Immunization (EPI) with the objective of reducing morbidity, mortality and disabilities from these six diseases by making free vaccination services easily available to all eligible children. The objective of full immunization coverage is a cornerstone of the Indian child health care system (International Institute for Population Sciences, 2000).

give information on such practical issues as the weekly immunization days in local health centers and block hospitals for example.

Apart from raising awareness among the participants, the trainings mobilize women to participate in the government Pulse Polio Campaign. Moreover, many women's groups initiate small immunization campaigns in their own village after having received a training. They go from door to door in their neighborhood to inform other families of the importance of immunization and about the free possibilities of immunization in the health centers.

Figure 1 shows the 2003 immunization rates of children aged 0 to 13 in programme and control villages. Within programme villages it also compares households participating in Mahila Samakhya with non-participating households. Since the control villages represent a good counterfactual, in the absence of externalities we would expect the immunization rates to be equal among non-participants and control households. However, the descriptive statistics show a substantially higher immunization rate for non-participants. This difference is both sizeable and statistically significant (see panel A of table 5).



Within programme villages, the difference between participants and non-participants is always positive but significant only for measles. On the one hand, this could indicate that the increased health knowledge spills over to a large extent to the rest of the community, thereby eliminating individual effects on participants. On the other hand, the participants usually belong to the least educated, poorest groups of a village and are thus expected to have a lower immunization rate. Indeed, the difference between the two groups becomes significant once we control for household characteristics. Panel B of table 5 indicates that overall, children belonging to the Scheduled Castes and Other Backward Castes are significantly less likely to be immunized against any of the four diseases (including polio) than General Castes children. Muslim children show an alarmingly low coverage rate compared to the total Hindu population, including the Scheduled Castes. Disaggregation by sex (panel C in table 5) does not show a significant difference between girls' and boys' immunization rates.

The government polio campaign seems very successful in reaching a large number of children although one out of every ten children is not reached. There is no significant difference with respect to polio vaccination between the programme or control villages, indicating that the government campaign is equally active in all blocks under study. Participation of the women's groups might facilitate the campaign but does not have a visible effect on immunization coverage. Therefore, polio is left out of most of the following analysis. Overall, immunization against tuberculosis is more prevalent than against DTP and measles. Apparently, the longer vaccination period has a

stronger positive effect than the potential downward effect of the longer travel distance. Immunization against measles is least common among participants, non-participants and control households alike.

4.2 Education

Improving educational outcomes is one of the focus issues of the Mahila Samakhya programme in Bihar. The programme aims to raise awareness among its members with respect to the importance of education, for girls as well as for boys. The facilitators and programme trainers also discuss with participants the role of preschool education in preparing young children for school.[18] Women's groups are stimulated and supported to set up informal preschools in their community (so-called bal jag jagi centers) as well as informal primary schools for girls who are not enrolled in formal school (jag jagi centers)[19].

The preschools aim to prepare young children aged 3 to 6 for entrance in primary school through games, stories and songs. By March 2002, the women's groups in Muzaffarpur had set up 137 preschools, in Sitamarhi they were operating 110 centers and in Darbhanga, the youngest programme district, 23 preschools had been opened (Mahila Samakhya, 2002).

For adolescent girls who have never been enrolled in primary school or have dropped out, the women's groups can open an informal girls' primary school in their village. These primary schools operate a few hours per day during which the girls receive skills-based education with a focus on daily life issues. A second objective of the girls' schools is mainstreaming. After passing three levels, a girl should be ready to enter the formal primary school system in the appropriate grade. In Muzaffarpur, 105 informal schools had been opened by March 2002, 115 in Sitamarhi and 74 in Darbhanga (Mahila Samakhya, 2002).

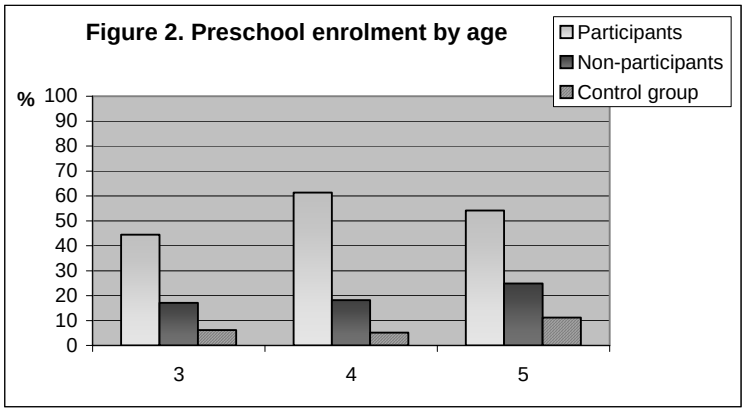
Both the awareness raising activities and the opportunity to set up informal pre- and primary schools are likely to result in higher enrolment. Table 6 shows descriptive statistics about preschools in the study area. On average 24.6% of the control villages and 66.9% of the programme villages have at least one preschool in their community[20]. Accordingly, the overall preschool enrolment rate in control villages is 7.7% compared to 22.2% in programme villages (53.4% among participants and

[18] For overviews of the evidence on the impact of early childhood education programmes, see for example Barnett (1995), Young (1997; 2002), Karoly et al. (1998) or Currie (2001). Individual impact evaluations of preschool programs are for example Behrman et al. (2004) for Bolivia, Paes de Barros and Mendonca (1999) in Brazil, Kagitçibasi et al. (2001) in Turkey, and the evaluation of Head Start in the United States (Garces et al., 2002).

[19] Both the preschools and the girls' primary schools are opened only on demand of the women's group. The Mahila Samakhya programme provides initial and follow-up training for the instructor and finances her honorarium. However, the operation of the centers is the responsibility of the group. The women need to find a (covered) space for the center and to maintain the center. They choose the instructor (usually a woman from their own village) and monitor her. Also, the group is responsible for the payments to the instructor.

[20] Not all preschools are financed by Mahila Samakhya. A number of preschools are Early Childhood Education centers run by the Bihar Education Project. The ICDS (Integrated Child Development Services) scheme of the Ministry of Women and Child Development operates so-called Anganwadi centers. The data do not allow us to distinguish between these types of preschools.

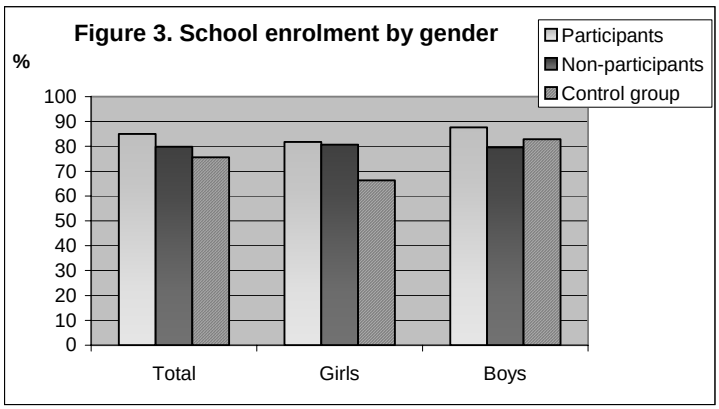
20.4% among non-participants). Table 6 provides more details. Figure 2 gives us an impression of preschool enrolment by age. It represents the percentage of children aged 3 to 5 years old that go to preschool. For each age group, we see that Mahila Samakhya members send their children to preschool considerably more often than non-participants from the programme villages. Also, non-participants from programme villages send their children significantly more often to preschool than parents from control villages. These differences are significant at the 1% error level. Again, if there were no externalities of the programme we would expect to see an enrolment rate for non-participants similar to that of children in the control group.



The higher preschool enrolment rate in programme villages might be related mainly to the increased availability of preschools in the community. If we consider only the villages with at least one preschool within the community boundaries the enrolment rates for control and program villages are 17.9% and 27.2% respectively (62.1% among participants and 25.1% among non-participants). Panel C of table 6 gives more details.

The third child outcome we consider is enrolment in school of children 6 to 13 years old. These children can either be enrolled in primary or middle school. The enrolment rates do not take into account whether a child is enrolled in the appropriate grade. Nor do we look at actual attendance, since unfortunately these data are not available. Table 7 gives an overview of the net school enrolment rates. School enrolment in the study area is 75.5% in control villages compared to 80.2% in programme villages. Within programme villages we find an enrolment rate of 85.0% for children from participating households and 79.9% for non-participants. Whereas participants are more likely to be enrolled than both non-participants and control children, the differences are not statistically significant.

Disaggregation of the data by gender shows an entirely different picture. Figure 3 provides a graphical representation of the total enrolment and the enrolment rates of girls and boys separately. Without controlling for household characteristics, boys from participating households are significantly more likely to be enrolled than boys from non-participating households.



Since members generally come from the poorest, least educated and socially marginalized groups, the descriptive statistics suggest a positive effect of the programme on the enrolment of boys whose mother participates. Differences between non-participating boys in programme villages and the control group are insignificant.

However, the results for girls show a large and highly significant difference between programme and control villages. Both members and non-members in programme villages are substantially more likely to enroll their daughters in primary or middle school. These results suggest that the programme not only affects girls' enrolment of participants, but has significant spillover effects on non-participants' daughters as well.

A similar pattern is revealed when comparing participants, non-participants and the control group by caste. Scheduled Castes households, whether participating in the programme themselves or not, shows a significantly higher enrolment compared to the control group. On the other hand for the Other Backward Castes we only find a significant increase in enrolment for participating households.

To summarize, the descriptive statistics on education suggest that the programme increases both awareness and access to preschools in programme villages. Participating households are more likely to send their child to preschool than other families in the community. In addition, children in programme villages are more likely to be enrolled in preschool, regardless of whether their own mother is a member of Mahila Samakhya. These findings are strongly suggestive of the presence of externalities. Regarding school enrolment, the effects of the programme seem more subtle. Although the results in this section do not control for a number of household and other characteristics, the tables and graphs suggest that the programme influences school enrolment in the participating households for both boys and girls. Externalities on other households in the community appear to exist mainly for girls and Scheduled Castes children.

5. Estimation methods

The descriptive statistics in the previous section give a first indication of the impact of the programme. Overall, members of a Mahila Samakhya women's group perform better on both immunization and school indicators as compared to non-participants. Likewise, children of households that are not participating themselves but that are living in a programme village are more likely to be immunized and to go to preschool. The primary and middle education spillovers seem limited to girls and Scheduled Castes children.

These results are merely indicative. First, participants usually belong to the more disadvantaged groups in their village in terms of caste, household education and income. Since these groups are also more likely to have low child outcomes, we will control for these household characteristics in order not to underestimate the impact of the programme. On the other hand, participation in the Mahila Samakhya programme is a voluntary decision. This could lead to a selection bias when comparing participants with non-participants. Participants might differ on additional unobservable characteristics such as awareness or motivation that could have a direct

effect on child outcomes beyond participation in the programme. That is, without the programme the immunization and enrolment rates of children of programme members might have been higher as well, due to unobserved characteristics.[21]

We will estimate the probability of immunization and enrolment using the following econometric specification:

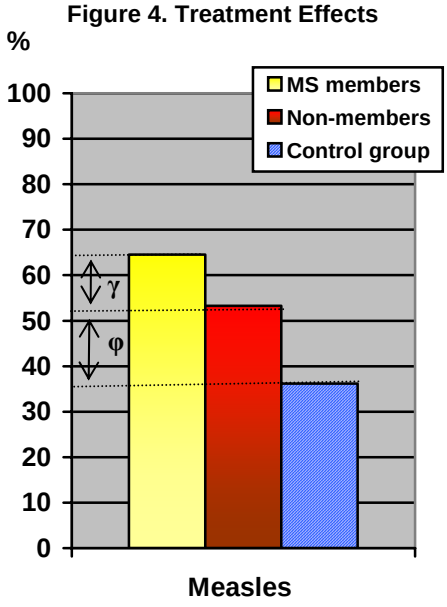
(1) $$Y_i = a + \beta X_i + \gamma P_i D_i + \phi P_i + \eta_i$$

where $Y_i = 1$ if child *i* is immunized / enrolled in (pre-) school; and 0 otherwise.

X_i is a vector of explanatory variables containing child, household and community characteristics. The child characteristics refer to the sex and age of the child. It includes a dummy for prior preschool enrolment in the estimation of primary/middle school enrolment. The household characteristics include dummies for Scheduled and Other Backward Castes as well as for belonging to the Muslim population. Moreover, they include a measure for household income, for the highest education level of adults in the household and the highest female education level in the household. Household size, a child dependency ratio and the gender of the head of household are included as well. The community characteristics refer to the distance to the nearest town, a village development indicator based on road quality and the availability of facilities, the 1991 (pre-programme) village level rate of female literacy and the 1991 village level percentage of Scheduled Castes population.[22] The immunization estimations include the distance to the nearest health center. The preschool and school enrolment estimations include the number of pre- and primary schools respectively, and the latter also includes the average child wage in the village. Moreover, we include block and district dummies to control for variation in unobserved regional characteristics.[23]

$P_i = 1$ if child *i* is living in a programme village; and 0 otherwise. $D_i = 1$ if a female household member of child *i* participates in the programme (i.e. in a women's group); and 0 otherwise. This specification allows us to estimate the effect of participation versus non-participation and of living in a programme village versus a control village. Figure 4 graphically shows these effects (not taking into account the remaining explanatory variables). The estimate for γ

Figure 4. Treatment Effects



[21] For a review of non-experimental methods for the evaluation of social programmes and the estimation of treatment effects, see for example Blundell and Costa Dias (2000).
[22] We have also included the number of members and the years of existence of the women's groups in estimations but these do not have a significant effect. Likewise, village population totals, higher order terms and interaction terms were included. These variables either have no significant effects or do not qualitatively change the results. Results are not shown here.
[23] In the school enrolment estimations, we use 1991 block variables on female literacy and 1991 % Scheduled Castes population instead of block dummies.

captures the effect of participating versus non-participating when living in a programme village. The estimate for ϕ represents the difference between living in a programme village and living in a control village (regardless of whether the individual belongs to a participating household or not). In other words, ϕ captures the spillover effects for non-participants. The sum of γ and ϕ is the treatment effect for participants in the programme compared to the control group.

To account for participation selection bias, we will use instruments for the decision to become a member. We use a two-stage instrumental variables approach with linear probability in both stages. The first stage reflects the participation decision and has the form:

$$(2) \qquad D_i = c + \lambda X_i + \delta Z_i + \mu_i$$

where X_i encompasses individual, household and community characteristics and Z_i contains the instruments, not included in X_i. This equation is estimated only for the population living in programme villages.[24] The estimated coefficients are used to predict participation for the entire population.[25] The predicted participation variable, $\hat{D}_i$, is entered as a new exogenous variable in the outcome equation and interacted with the programme village dummy:

$$(3) \qquad Y_i = a + \beta X_i + \gamma P_i \hat{D}_i + \phi P_i + \eta_i$$

Since we assume that there is no village selection bias, the estimate for ϕ captures the spillover effect, or –in case of village treatment heterogeneity– the spillover effect on the treated (i.e. on the programme villages). Although programme villages might have additional unobserved gains from being in the programme as compared to control villages, it does not seem likely that there is systematic village treatment heterogeneity after controlling for the block selection criteria and village variables. Appendix A gives a more detailed analysis of the treatment effects estimated when using the above specification.

In the absence of participation treatment heterogeneity, the estimate for γ reflects the average treatment effect of participation. The sum of the estimates $\hat{\gamma}$ and $\hat{\delta}$ reflect the total treatment effect for participants compared to the control group.

However, when there is participation treatment heterogeneity due to unobserved characteristics of participants, the instrumental variables estimate $\hat{\gamma}$ of the average treatment effect on the treated is likely to be biased (Heckman, 1997).[26] This will be

[24] Appropriate weights in the participation equation correct for choice-based sampling of participants and non-participants in programme villages (Manski and Lerman, 1977).
[25] For more details on two-sample two-stage instrumental variables estimation, see Angrist and Krueger (1992).
[26] When using instrumental variables if treatment heterogeneity is present, the estimated treatment effect represents a local average treatment effect (LATE), or the effect of the programme on those individuals that change their decision to participate with a change in the instruments (Imbens and Angrist, 1994; Angrist et al., 1996). Such an interpretation makes sense when the instrument denotes different policy regimes. However, in our case the instruments refer to personal attitudes and characteristics, and the interpretation of the LATE is not straightforward.

the case when individuals act on private information about their personal gains from the programme relative to others that cannot fully be predicted by observable variables.

It is not clear whether participants in the Mahila Samakhya programme have additional unobserved relative gains from participation as compared to non-participants. On the one hand, it seems likely that especially those women who expect to gain most will participate. On the other hand, some evidence suggests otherwise. When asked to list the benefits of participation, both participants and the non-participants who know about Mahila Samakhya give a highly similar response.[27] Likewise, a reason for non-participants to expect *less* relative unobserved gains could stem from conflicts within the household because of participation. But again, the percentage of non-participants that would expect domestic conflicts from participation is not significantly different from the percentage of participants that actually experienced problems.[28] Finally, it might be possible that women do not know beforehand what will be their personal gains from participation relative to others. Then the relative individual gains in child outcomes (due to unobservable characteristics) will not influence their participation decision in a systematic way.

In sections 6.1 to 6.3 we will estimate the treatment effects using instrumental variables. In doing so we assume that participation treatment heterogeneity is negligible. In section 6.4, we will again estimate the treatment effects of the programme using propensity score matching instead. This allows for participation treatment heterogeneity, but it assumes on the contrary that participation selection bias on unobservables is negligible.

6. Estimation of the treatment effects

6.1 Identification of the instrumental variables

To correct for the potential participation selection bias, we need instrumental variables that are correlated with participation in the programme but not correlated with the child outcomes. It is usually a difficult task to find these instruments. We propose a number of variables that are related to programme participation. To test whether they can be excluded from the outcome equations, we estimate the likelihood

[27] We asked both participants and non-participants to name the most important benefits of participation in Mahila Samakhya. About 45% of the non-participants do not know enough about Mahila Samakhya to respond to these questions. The participants and the non-participants who know the programme mention the following benefits (as a percentage of all benefits they mention and in decreasing order). Increased literacy: 27% vs 29% (participants vs. non-participants), increased health knowledge: 23% vs. 25%, access to credit group: 18% vs. 20%, emotional support: 13% vs. 6%, improved education for children: 6% vs. 9%, empower as a woman within the community: 7% vs. 2%, access to government subsidies: 3% vs. 5%, improved status within the household: 2% vs. 3%, no benefits at all: 1% in both groups. Better access to water, improved infrastructure and 'other benefits' where mentioned in very few cases. The differences between participants and non-participants are highly insignificant except for the benefits derived from empowerment within the community and emotional support. (Results are weighted for the survey design and the standard errors are adjusted accordingly).

[28] Of the participants, 16% indicate conflicts with their husband or family-in-law because of less time for domestic chores due to participation and 19% indicate problems within the household because of organizing as a woman. For the non-participants that are familiar enough with the programme to answer the question, 17% and 22% can imagine to have these respective conflicts. The differences between participants and non-participants are insignificant.

of immunization and enrolment for control villages only. Apart from the child, household and community characteristics, these estimations contain the potential instruments. Since the programme is not active in these villages, we would expect to find a significant coefficient for the instruments if they are directly related to the child outcomes. If in contrast the coefficients of the instruments are small and statistically insignificant, this provides us with an argument to exclude the instrumental variables from the outcome equation. In that case, they will be used as instruments in the following sections. A test of overidentifying restrictions gives additional support to the exclusion restriction in those cases where two instrumental variables are retained.

Two indicators are identified as potential instruments. The indicators are scores derived from a factor analysis of a number of statements. In particular, we consider a 'civicness'-indicator and an 'assistance'-indicator. [29]

Table 8 describes the potential instruments (panel A), compares their means for participants, non-participants and control households (panel B) and shows the correlation matrix of participation and instruments (panel C). Although the instruments show sufficient variation and cover the same range for the three sample groups, their (weighted) mean varies among the three groups. On average participants in the Mahila Samakhya programme are more likely to have engaged in civic behavior during the last three years than non-participants. They also give more unpaid assistance to persons outside the household. Non-participants display on average slightly less civic behavior and give substantially less often assistance to others. Only the latter difference is statistically significant. The correlation matrix in panel C of table 8 uses only observations from programme villages. Especially 'civicness' is highly correlated with participation in Mahila Samakhya.

The first condition to be satisfied is the nontrivial correlation with the participation variable. To further investigate this condition, table 9 shows the results of a probit estimation of the likelihood of participation in Mahila Samakhya. The observations are weighted for outcome-based sampling. Standard errors are robust and corrected for clustering at the village level. The control households are omitted from this estimation, since they do not have the opportunity to participate. Both potential instruments are significantly correlated with participation. These results are robust to changes in indicator calculations (using factor analysis, principal components or equal weights), and changes in the specification (including village and block dummies, block variables, additional trust and cooperation related variables, and higher order terms).

The second requirement for the validity of instrumental variables is the exclusion restriction. Instruments should not directly influence the child outcomes, only indirectly through their relation with participation. This requirement is often

[29] The two indicators are calculated using the loadings of the first rotated factor of a factor analysis.

- The 'civicness'-indicator is based on four types of civic behavior based on the answers to the questions: "In the last three years have you personally: a) Voted in the elections?, b) Contacted your elected representative about a community problem? c) Taken part in a protest march, demonstration or sit-in? and d) Actively participated in an information campaign?".
- The 'assistance'-indicator is based on the answers to the following questions: "In the past year, has anyone in the household assisted someone else outside the household with: a) looking after children without getting paid; b) preparing food without getting paid; c) helping to build a house without getting paid; d) working without getting paid; or e) giving financial aid?".

problematic as it is difficult, in general, to find such variables and to argue convincingly that the exclusion restriction is satisfied.

Our sample design allows us to take a different approach. We will use the control group to test whether the potential instruments are partially correlated with the child outcomes. Since the programme is not active in these villages, any direct effect will be captured in a significant partial correlation of the instrumental variables with the outcomes. Table 10 gives the results of the probit estimations for each of the child outcomes under study. The usual child, household and community characteristics are included as explanatory variables.

As the results show, both the 'civicness' and the 'assistance' indicators are not significantly related to any immunization outcome. A test of joint insignificance of potential instruments, given in the last row of the table, is not rejected. Changes in the specification (e.g. different calculation methods for the indicators, linear probability instead of probit, including and excluding village and block variables, introducing higher order and interaction terms) test the robustness of these results. Since the assistance-index is not robust for changes in the estimation of tuberculosis, it will not be used as an instrument for that type of immunization. Although 'civicness' is slightly significant for both preschool and school enrolment (but not at the 5% level), a test of joint insignificance of the two instruments is not rejected. Nonetheless, these estimations suggest that the instrumented results for education should be treated with some caution.

The final instruments to be retained for each outcome are listed in table 11. The last row of the table gives the χ^2-values of a test of overidentifying restrictions for those child outcomes with more than one potential instrument. A test of overidentifying restrictions does not reject the validity of the instruments at the 5% error level. We will retain 'civicness' and 'assistance' as an instrument for all outcomes, except for tuberculosis for which we will only use the former variable.

6.2 Estimation of immunization

Table 12 shows probit estimations of the likelihood of immunization for children aged 0 to 13 years for the four types of vaccinations. Note that participation is not yet instrumented in these estimations. These first results clearly show the difference between polio on the one hand, and tuberculosis, DTP and measles on the other hand. Whereas the participation and the programme village dummy are highly significant for the latter three immunization probabilities, immunization against polio seems to be unrelated to the presence of Mahila Samakhya. Moreover, a number of household characteristics that are significantly correlated with the likelihood of immunization against tuberculosis, DTP and measles, such as household education, female education, household size and dependency ratio, are unrelated to the polio immunization rate. However, this is not surprising given the massive, nationwide nature of the polio campaign. An analysis of the mechanisms underlying polio immunization falls outside the scope of this paper.

After the identification of the instrumental variables, we can estimate the effect of the programme on immunization using the two-stage instrumental variables approach described in section 5. Tables 13 to 15 give us the two-stage instrumental variables

(2S IV) estimations of the three immunization probabilities for the total child population and for several subgroups of children. As table 13 shows, the results for tuberculosis do not qualitatively differ when using a probit estimation (first column), a probit estimation with marginal probabilities[30] (second column), linear probability without IV (third column) and the two-stage least squares estimation with IV (fourth column). In all four cases, the coefficients for participation and programme village are highly significant at the five or one percent error level. We correct the standard errors for clustering at the village level. Also, the standard errors for the two stage estimation are bootstrapped with two hundred replications. The coefficients are of similar magnitude except the 2S IV estimates that are considerably larger. This suggests that participants may indeed belong to the most uninformed and marginalized women, whom the programme intends to target.

Since the effect of the programme for participants versus non-participants in programme villages is captured in the participation variable, the programme village dummy reflects the impact of living in a programme village versus control village *regardless of whether the household itself is participating.* In other words, the coefficient for the programme village dummy measures the spillover effects for non-participating households. This same coefficient combined with the participation coefficient measures the effect for participants relative to the control group.

The coefficients of the control variables have the expected sign[31] and are generally highly significant. In the first four specifications (columns (i) to (iv)) we find a strong negative coefficient for girls and for age. The latter indicates that the immunization coverage rate is increasing over time, with children who were born more recently also more likely to be immunized. The only difference between the instrumented and the non-instrumented specification is the negative coefficient for Scheduled Castes and Other Backward Castes, which is highly significant only in the instrumented version. Household education and income have a positive relation with immunization. The significant coefficient for female education indicates a strong and additional effect of the education of women. Larger households are less likely to immunize their child. This is also true for households with more children per adult. Distance to the nearest town and distance to the nearest health center are not significant for tuberculosis immunization. The estimations remain largely unchanged when we consider only girls or only boys, although girls' immunization appears more dependent on the distance to the health facility. When we consider only Scheduled Castes children, the programme village variable becomes insignificant although it remains large and positive. Note that an interaction term between programme village and Scheduled Castes in the full sample is significant and positive, which would suggest increased impact on the lowest caste children.

The results for the likelihood of immunization against DTP (table 14) and measles (table 15) are roughly the same. Again, we find a strong and positive impact for

[30] The marginal probabilities in column (ii) of table 13 are based on calculations for the 'average' child. Since most of the variables used in the estimations are dummies, we also calculate the marginal effect of the programme based on the average of *individual* marginal effects. See table 18 for the results.
[31] For example, Partha and Bhattacharya (2002) estimate the probability of immunization in four Indian states including Bihar. They find odds ratios below one for girls, and for increasing birth orders, for Scheduled Castes children and Muslim children. Odds ratios above one were found for increasing levels of father's education, and for mother's education in addition to father's education.

participation in the programme as well as for living in a programme village. An interesting difference with tuberculosis is the finding that the distance to the nearest health center becomes significant. Whereas tuberculosis is administered in block hospitals (and hence, distance to nearest town would be expected to play a role), DTP and measles vaccinations are given in the health centers. Indeed, children living further away from a health center are less likely to have received the vaccination against these two diseases.

In summary, women who participate in a Mahila Samakhya women's group are significantly more likely to immunize their child against tuberculosis, DTP, measles and polio than women who live in a programme village but who do not participate themselves. In addition, the highly significant coefficient for programme village indicates that non-participating women living in programme villages are also significantly more likely to immunize their child against tuberculosis, DTP and measles (but not polio) than women living in a village where the programme is not active. That is, the spillover effects on immunization coverage within programme villages is substantial. The only exception is immunization against polio.

6.3 Estimation for educational outcomes

Table 16 shows the results of several estimations for preschool enrolment of children aged 3 to 5 years. The first column shows the results of a simple linear probability estimation when we include all programme and control villages. As expected, the coefficient for the number of preschools within the village is highly significant.[32] This coefficient reflects the importance of access in order to increase preschool enrolment. Simply put, if there is no preschool present in the village it becomes very difficult for parents to enroll their child. As for immunization coverage, we find large and significant effects both for participants versus non-participants, and for non-participants versus the control group. Again, these findings suggest that the programme has substantial externalities for the wider community.

In order to separate to some extent the effect of increased access from increased awareness, the remaining columns of table 16 show estimations of the likelihood of preschool enrolment *only* for those villages with at least one preschool. For the total population as well as for girls and boys separately the village dummy is very large and highly significant. Since access is not the main problem in these villages, it appears that the programme has been able to raise awareness about the importance of early childhood education[33]. Parents with equal possibilities send their children to preschool more often when they live in a Mahila Samakhya village compared to control villages, regardless of their own participation in the programme. The non-instrumented and instrumented participation estimates are of equal size.

Most of the child and household characteristics included in the estimations are not significant. A possible explanation for the absence of significant effects of these characteristics might be that the influence of the Mahila Samakhya programme is

[32] Standard errors are robust and corrected for clustering at the village level. In the 2S IV estimation, standard errors are also bootstrapped with 200 replications.

[33] The introduction of village population size and an interaction term of village population size with the number of preschools in a village does not change the results. Population size and the interaction term are jointly and individually insignificant (with a negative coefficient for population size).

much stronger than the effect of other household characteristics. Hence, the usual household indicators might become insignificant. Children living in households with a higher child dependency ratio are more likely to be enrolled, suggesting that the need for child care is largest among households with more children per adult. Access plays an important role as well. Preschool enrolment is positively related to additional preschools in a community. Note that, conditional on there being at least one preschool, households in the least developed villages are most likely to enroll their child.

Finally, we measure the impact of the programme on primary and middle school enrolment. Table 17 shows the results for enrolment in primary and middle school of children in the age group of 6 to 13 years. We find a highly significant positive effect of own participation in the programme, both in the instrumented and non-instrumented estimations. The programme village dummy, which captures the spillover effects, is not significant for the total sample or for boys. It is significant and substantial for the subsample of girls and of Scheduled Castes children. However, these results hold only if we include an interaction term between programme village and 1991 pre-programme village female literacy levels. Whereas these two variables individually are positive and significant, their interaction term is negative. This indicates that the programme indeed has spillover effects on non-participant girls (and SC children) in the community, but mainly in those villages where pre-programme female literacy rates were low. In higher educated villages, the programme externalities are smaller and statistically insignificant.

Again, the estimates for the control variables do not change much between the different estimation methods. They all have the expected sign[34] and significance[35]. Girls and Scheduled Castes children or Muslim children are less likely to be enrolled in school. Household education and income have a positive effect. Female education in the household has an additional positive and significant coefficient for girls' enrolment only. Children living in larger households with more children are less likely to go to school. The positive coefficient for age signals late enrolment (age-squared has a negative and significant coefficient, indicating that this effect tops off). Finally, children that went to preschool at a younger age are more likely to be enrolled in primary / middle school. This could capture a higher awareness of parents regarding education. It could also stem in part from the stimulating effects of preschool on school preparedness and learning achievement.

[34] Since primary education is not compulsory in Bihar, a household's decision to send a child to school will depend both on the costs of education and on the perceived benefits. Costs relate for example to tuition fees (absent or very low in the study region), travel costs, costs of books and uniforms, and opportunity costs of foregone child labor. The benefits are longer term and concern not only future income but also a better ability of women to take care of their family for example [Glewwe, 2002 #158]. To illustrate, a higher education of Indian mothers is found to be significantly related to reduced incidence of diarrhoea among children (Borooah, 2004; Jalan and Ravallion, 2003), to reduced fertility (Drèze and Murthi, 2001), and to school enrolment of their children (Behrman et al., 1999), especially their daughters [Pal, 2004 #140]. Whereas the costs to the family are immediate, the potential benefits of education are not always apparent. For girls in rural Bihar moreover, these benefits will usually not accrue to her own family. Upon marriage, women are mostly expected to move and live with their parents-in-law [Nabar, 1995 #153].

[35] Standard errors are robust and corrected for clustering at the village level. In the 2S IV estimation, standard errors are also bootstrapped with 200 replications.

6.4 Propensity score matching

To test our results further, we also use the propensity score matching (PSM) method to estimate the average treatment effects on the treated for participants versus non-participants, and the spillover effects on non-participants versus the control group. This method is explicitly based on the assumption that, given the observable covariates, participation in the program is independent of the outcomes. That is, observable variables capture any systematic differences between participants and non-participants. Other than those differences, there would be no systematic difference in the outcome if the program had not been implemented (i.e. we assume that $E(Y^0 \mid X, D=0) = E(Y^0 \mid X, D=1)$).

While the comparability of the treatment and control group increases with an increasing number of covariates, a larger number of controlling variables also significantly adds to the complexity of matching the two groups. However, Rosenbaum and Rubin (1983) show that whenever it is valid to match on all the covariates separately, it is also valid to match on the propensity score. The propensity score is the conditional probability of participation given a vector of observed covariates.

We estimate the propensity score (the probability of participation in the programme) within programme villages only. The likelihood of participation is estimated with a probit model, using appropriate weights to correct for the choice-based sample selection procedure within programme villages. The covariates consist of the child, household and community variables as described in section 5 (except the instrumental variables).[36] Using the coefficients from this estimation, the propensity scores were calculated for the participants, the non-participants and the control group. A regression of the participation dummy on the covariates and the propensity score provides a first test of the balancing requirement. Except for the coefficient of the propensity score, none of the coefficients are significant, either jointly or individually.[37]

Matching occurred over the common support only, using nearest neighbor matching with replacement. To estimate the effects of participation versus non-participation when living in a programme village, participants are matched with non-participants. A test of the balancing requirement[38] confirms that after matching, the differences between the sample means (as a percentage of the square root of the average sample variance) for each covariate in the two sub-groups are not significant.

The total direct effect and the spillover effect are estimated by matching participants and non-participants respectively with children from comparable households in the

[36] Instead of the calculated income quintile and village development index, both derived from factor analysis, we use the underlying individual variables for more detailed matching.
[37] Results available upon request.
[38] The balancing requirement is tested using the pstest-command in Stata suggested by Sianesi and Leuven (2001), http://ideas.repec.org/c/boc/bocode/s432001.html . This tests calculates the reduction in covariate imbalance after propensity score matching for each covariate. The imbalance or bias is the difference of the sample means in the treated and non-treated sub-samples as a percentage of the square root of the average of the sample variances in the treated and non-treated group.

control group. Thus, we estimate the propensity score as the likelihood of participation also for the individuals in the control group, and use this estimate to match children in programme villages with similar children in control villages. Using 20 as a cut-off point, some of the distance variables, especially distance to town, do not balance.[39] However, their exclusion from the propensity score does not significantly change the results. The only exception is the estimated participation impact on boys' school enrolment which is substantially smaller using the restricted specification, but still positive and statistically significant at the 10 percent level.[40]

The results from PSM confirm the presence of both the direct impact and the spillover effect of the programme. Table 18 provides a comparison of the results from five alternative estimation methods. Consider first the results for immunization. The estimated treatment effects on 'participants' versus 'non-participants' when living in a programme village are given in columns (i), (ii) and (v). The participation effects are very similar among all methods and indicate a direct participation effect between 8 and 10 percentage points. As discussed in the previous section, the instrumental variables estimates are substantially larger.

The village dummy in columns (i) to (iii) captures the spillover effect of the programme on non-participants versus the total population in control villages. Column (iv) gives an estimate of the impact on the village as a whole. The estimates are of the same order of magnitude. The average spillover effect derived from PSM (column vi) is considerably lower than the estimates found with the other methods. However, the comparison group used in column (vi) is a different one: non-participants are matched with control children with similar propensity scores instead of the entire control group. Since non-participants are relatively better off, their matched counterparts are also more likely to be immunized than the average for all children in control villages. The village effects remain very substantial at a 9 to 11 percentage points' increase in immunization coverage for non-participants. The results for polio consistently indicate the absence of treatment effects.

The total direct impact of the programme on participants is captured by the sum of the participation and village effect in columns (i) to (iii). The PSM counterpart of this total impact is given in column (vi). The impact estimated from PSM (where participants are matched with control children showing comparable propensity scores) is considerably smaller. Since children of (would-be) participants are generally worse off than the average child is, we would expect the total impact estimated with PSM to be higher. The difference might be due to the different support. Other recent findings also suggest that PSM estimates do not precisely replicate experimental or regression results (Agodini and Dynarski, 2004; Smith and Todd, 2004).

The second panel of table 18 provides a comparison between the various methods for preschool enrolment in villages with at least one preschool. Except for PSM, all methods estimate the effect on preschool enrolment of participation versus non-participation at 21 percent and the village or spillover effect between 34 and 44 percent. Based on the propensity score matching method (column vi) the total direct

[39] For a few sub-samples the Muslim indicator or the dependency ratio does not balance either.
[40] The impact estimate using the restricted propensity score estimation is .066 with a standard error of .039.

effect on participants is lower but substantial at 34 percent and the total spillover effects on non-participants is 19 percent.

The last panel of table 18 looks at the impact of the programme on school enrolment. Participation estimates for the total sample as well as for girls and boys separately indicate that participants are substantially more likely than non-participants and the control group to enroll their child. The total effect is estimated to be 11 percent using the PSM results. With spillover effects of at least 14 percentage points, externalities on girls living in programme villages are estimated to be almost as large as the direct effect. That is, once a Mahila Samakhya group is active in a village, girls' school enrolment seems to increase regardless of whether her mother participates. Such externalities appear to be absent for boys.

Using these results, it is possible to calculate the size of the spillover effect on non-participants as a percentage of the total direct effect for participants. We use the treatment effects estimated with the propensity score method in column (vi). These are the most conservative estimates among the alternative methods. The estimated spillover effects show to be a very large percentage of the participation effects. In the case of immunization, the programme externalities on non-participants are at least forty percent of the direct programme effect on participants (excluding the insignificant estimates for polio): 54 percent for tuberculosis, 44 percent for DTP and 40 percent for measles. Likewise, we find a spillover effect on preschool enrolment of non-participants that is equal to 54 percent of the size of the direct programme effects. Finally, programme spillovers on school enrolment of non-participants are 49 percent of the total impact on participants, but they are not significant for the sub-sample of boys.

7. Conclusion

The empowerment of low-caste low-educated women in India can have far-reaching consequences not only for the women themselves, but also for their children and for the wider community in which they live. An impact evaluation of the Mahila Samakhya programme in northern Bihar that specifically aims at those most marginalized women, shows that the mobilization of women into women's community groups can generate large externalities.

The programme has a significant impact on the children of women participating in the programme. The participants are considerably more likely to immunize their children, and to send them to preschool and school than are parents who do no participate in the programme. Moreover, we find significant spillover effects of the programme on households that live in villages where a women's group is active, but that do not participate in the programme themselves. These externalities are especially strong for immunization against tuberculosis, DTP and measles (but not for polio), for enrolment in preschool, and for enrolment in school of girls (but not of boys).

We find these treatment effects regardless of the method we use to estimate impact. Simple probit, linear probability, two stage instrumental variables, and propensity score matching yield results of comparable magnitude, although PSM estimates are generally lower on average. Nonetheless, the externalities remain significant and

large. They are equal to at least forty percent of the direct programme impact on participants' immunization coverage, and half of the direct programme impact on preschool and school enrolment.

These results show the importance of using the right comparison groups in programme evaluation to avoid a potentially double underestimation of the programme effects. Especially when an empowerment or community programme emphasizes awareness raising and information dissemination, combined with a focus on collective action, these externalities can be substantial. The design of the survey is crucial in measuring these effects. Moreover, the inclusion in the sample of a control group where the programme is not yet active allows for a direct test of the exclusion restriction needed to correct for participation selection bias. Especially programmes that are still at the beginning of a scaling-up process might offer great opportunities for quasi-experimental survey designs.

References

Agarwal, Bina (2001), "Participatory Exclusions, Community Forestry, and Gender: An Analysis for South Asia and a Conceptual Framework", *World Development* **29**(10): 1623-1648.

Agodini, Roberto and Mark Dynarski (2004), "Are Experiments the Only Option? A Look at Dropout Prevention Programs", *Review of Economics and Statistics* **86**(1): 180-194.

Alderman, Harold, Jooseop Kim and Peter Orazem (2003), "Design, Evaluation, and Sustainability of Private Schools for the Poor: The Pakistan Urban and Rural Fellowship School Experiments", E*conomics of Education Review* **22**: 265-274.

Angrist, Joshua D., Guido W. Imbens and Donald B. Rubin (1996), "Identification of Causal Effects Using Instrumental Variables", *Journal of the American Statistical Association* **91**(434): 444-455.

Angrist, Joshua D. and Alan B. Krueger (1992), "The Effect of Age at School Entry on Educational Attainment: An Application of Instrumental Variables With Moments From Two Samples", *Journal of the American Statistical Association*, **87**(418): 328-336.

Barnett, Steven (1995), "Long-Term Effects of Early Childhood Programs on Cognitive and School Outcomes", *The Future of Children* **5**(3): 25-50.

Behrman, Jere R., Yingmei Cheng and Petra E. Todd (2004), "Evaluating Preschool Programs when Length of Exposure to the Program Varies: A Nonparametric Approach", *Review of Economics and Statistics* **86**(1): 108-132.

Behrman, Jere R. et al. (1999), "Women' Schooling, Home Teaching, and Economic Growth", *Journal of Political Economy* **107**(4): 682-714.

Besley, Timothy and Anne Case (1994), "Diffusion as a Learning Process: Evidence from HYV Cotton", Discussion Paper 174, May 1994, Princeton, NJ: Princeton University.

Blundell, Richard and Monica Costa Dias (2000), "Evaluation Methods for Non-Experimental Data", *Fiscal Studies* **21**(4): 427-468.

Bobonis, Gustavo J. and Frederico S. Finan (2002), "Do Transfers to the Poor Increase the Schooling of the Non-Poor: The Case of Mexico's Progresa Program", mimeo, December 2002, Berkeley: University of California.

Borooah, Vani K. (2004), "On the Incidence of Diarrhoea Among Young Indian Children", *Economics and Human Biology* **2**: 119-138.

Case, Anne C. and Lawrence F. Katz (1991), "The Company You Keep: The Effects of Family and Neighborhood on Disadvantaged Youths", NBER Working Paper 3705, Cambridge, MA: National Bureau of Economic Research.

Currie, Janet (2001), "Early Childhood Education Programs", *Journal of Economic Perspectives* **15**(2): 213-238.

Das Gupta, Monica (1987), "Selective Discrimination against Female Children in Rural Punjab, India", *Population and Development Review* **13**(1): 77-100.

Drèze, Jean and Mamta Murthi (2001), "Fertility, Education, and Development: Evidence from India", *Population and Development Review* **27**(1): 33-63.

Drèze, Jean and Amartya Sen (2002), *India: Development and Participation*, Oxford: Oxford University Press.

Duflo, Esther (2003), "Scaling up and Evaluation", ABCDE Conference, Bangalore.

Duflo, Esther and Emmanuel Saez (2002), "The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment", Social Science Research Network Paper 02-23, March 21, 2002, Cambridge, MA: MIT Department of Economics.

Filmer, Deon, Elizabeth M. King and Lant Pritchett (1998), "Gender Disparity in South Asia: Comparisons Between and Within Countries", Policy Research Working Paper WPS1867, January 1998, Washington, DC: World Bank.

Filmer, Deon and Lant Pritchett (2001), "Estimating Wealth Effects without Expenditure Data -- or Tears: An Application to Educational Enrollments in States of India", *Demography* **38**(1): 115-132.

Foster, Andrew D. and Mark R. Rosenzweig (1995), "Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture", *Journal of Political Economy* **103**(6): 1176-1209.

Garces, Eliana, Duncan Thomas and Janet Currie (2002), "Longer-Term Effects of Head Start", *American Economic Review* **92**(4): 999-1012.

Glewwe, Paul (2002), "Schools and Skills in Developing Countries: Education Policies and

Socioeconomic Outcomes", *Journal of Economic Literature* **40**: 436-482.
Government of India (2002), *National Human Development Report 2001*, New Delhi: Planning Commission, Government of India.
Heckman, James J. (1997), "Instrumental Variables: A Study of Implicit Behavioral Assumptions Used in Making Program Evaluations", *Journal of Human Resources* **32**(3): 441-462.
Heckman, James J., Robert J. Lalonde and Jeffrey A. Smith (1999), "The Economics and Econometrics of Active Labor Market Programs", *Handbook of Labor Economics*, A. Ashenfelter and D. Card, Amsterdam: Elsevier Science B.V. **3:** 1865-2097.
Hoff, Karla and Priyanka Pandey (2004), "Belief Systems and Durable Inequalities: An Experimental Investigation of Indian Caste", research paper (forthcoming), April 2004, World Bank: Washington.
Imbens, Guido W. and Joshua D. Angrist (1994), "Identification and Estimation of Local Average Treatment Effects", *Econometrica* **62**(2): 467-475.
India Census Office (2001a), "Census of India 2001: Bihar. Provisional Population Totals 2: Rural-urban Distribution of Population." Series 11, Patna, Bihar: Census Operations.
India Census Office (2001b), "Census of India 2001: Provisional Population Totals. Paper 1 of 2001" Series 1, New Delhi: Census of India.
International Institute for Population Sciences (2000), *National Family Health Survey (NFHS-2) 1998-99: India (Chapter 6: Mortality, Morbidity, and Immunization)*, Mumbai: IIPS.
Jalan, Jyotsna and Martin Ravallion (2003), "Does Piped Water Reduce Diarrhea for Children in Rural India", *Journal of Econometrics* **112**: 153-173.
Kagitçibasi, Çigdem, Diane Sunar and Sevda Bekman (2001), "Long-Term Effects of Early Intervention: Turkish Low-Income Mothers and Children", *Journal of Applied Developmental Psychology* **22**: 333-361.
Karoly, Lynn et al. (1998), "Investing in our Children: What We Know and Don't Know About the Costs and Benefits of Early Childhood Interventions", Discussion paper, Santa Monica: RAND Corporation.
Katz, Lawrence F., Jeffrey R. Kling and Jeffrey B. Liebman (2001), "Moving to Opportunity in Boston: Early Results of a Randomized Mobility Experiment", *Quarterly Journal of Economics* **116**(2): 607-654.
Kim, Jooseop, Harold Alderman and Peter Orazem (1999), "Can Private School Subsidies Increase Enrollment for the Poor? The Quetta Urban Fellowship Program", *World Bank Economic Review* **13**(3): 446-465.
Kishor, Sunita (1993), ""May God Give Sons to All": Gender and Child Mortality in India", *American Sociological Review* **58**(2): 247-265.
Lahiri-Dutt, Kuntala and Gopa Samanta (2002), "State Initiatives for the Empowerment of Women of Rural Communities: Experiences from Eastern India", *Community Development Journal* **37**(2): 137-156.
Lalive, Rafael and Alejandra Cattaneo (2004), "Social Interactions and Schooling Decisions", mimeo, October 2004, Zurich: University of Zurich.
Mahila Samakhya, Bihar (1995), "Mahila Samakhya Bihar: A Journey in Women Empowerment", report, Patna: Bihar Education Project.
Mahila Samakhya, Bihar (2002), "Mahila Samakhya Bihar: Forward & Onward 1992-2002", report, Patna: Mahila Samakhya Bihar.
Manski, Charles F. (1993), "Identification of Endogenous Social Effects: The Reflection Problem", *Review of Economic Studies* **60**(3): 531-542.
Manski, Charles F. (2000), "Economic Analysis of Social Interactions", *Journal of Economic Perspectives* **14**(3): 115-136.
Manski, Charles F. and Steven R. Lerman (1977), "The Estimation of Choice Probabilities from Choice-Based Samples", *Econometrica* **45**(8): 1977-1988.
Mansuri, Ghazala and Vijayendra Rao (2004), "Community-Based and -Driven Development: A Critical Review", *World Bank Research Observer* **19**(1): 1-39.
Miguel, Edward and Michael Kremer (2004), "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities", *Econometrica* **72**(1): 159-217.
Munshi, Kaivan (2004), "Social Learning in a Heterogeneous Population: Technology Diffusion in the Indian Green Revolution", *Journal of Development Economics* **73**: 185-213.
Munshi, Kaivan and Jacques Myaux (2002), "Development as a Process of Social Change: An Application to the Fertility Transition", mimeo, December 2002, Pennsylvania: University of Pennsylvania.
Nabar, Vrinda (1995), *Caste as Woman*, New Delhi: Penguin Books India.

Paes de Barros, Ricardo and Rosane Mendonca (1999), "Costs and Benefits of Preschool Education in Brazil", in collaboration with the ECD Team, World Bank, November 1999, Rio de Janeiro: Institute of Applied Economic Research.
Pal, Sarmistha (2004), "How Much of the Gender Difference in Child School Enrolment can be Explained? Evidence from Rural India", *Bulletin of Economic Research* **56**(2): 133-158.
Pande, Rohini P. (2003), "Selective Gender Differences in Childhood Nutrition and Immunization in Rural India: The Role of Siblings", *Demography* **40**(3): 395-418.
Partha, De and B.N. Bhattacharya (2002), "Determinants of Child Immunization in Four Less-Developed States of North India", *Journal of Child Health Care* **6**(1): 34-50.
Patel, Ila (1998), "The Contemporary Women's Movement and Women's Education in India", *International Review of Education* **44**(2-3): 155-175.
Purushothaman, Sangeetha (1998), *The Empowerment of Women in India: Grassroots Women's Networks and the State*, New Delhi: Sage Publications.
Ramachandran, Vimala (1998), "Bridging the Gap Between Intention and Action: Girls and Women's Education in South-Asia. The Indian Experience", New Delhi: UNESCO & ASPBAE.
Rao, Nirmala, Kai-Ming Cheng and Kirti Narain (2003), "Primary Schooling in China and India: Understanding How Socio-Contextual Factors Moderate the Role of the State", *International Review of Education* **49**(1-2): 153-176.
Rosenbaum, Paul R. and Donald B. Rubin (1983), "The Central Role of the Propensity Score in Observational Studies for Causal Effects", *Biometrika* **70**(1): 41-55.
Sacerdote, Bruce (2001), "Peer Effects with Random Assignment: Results for Dartmouth Roommates", *Quarterly Journal of Economics* **116**(2): 681-704.
Smith, Jeffrey A. and Petra E. Todd (2004), "Does Matching Overcome LaLonde's Critique of Nonexperimental Estimators?", *Journal of Econometrics*, **125**: 305-353.
Thorat, Sukhadeo (2002), "Oppression and Denial: Dalit Discrimination in the 1990s", *Economic and Political Weekly*: 572-578.
World Bank (1997), "India -- Third District Primary Education Project Bihar (DPEP III)", Project Appraisal Document 17102, October 27, 1997, Washington, DC: World Bank.
World Bank (2001), *Engendering Development: Through Gender Equality in Rights, Resources, and Voice*, New York: Oxford University Press.
Young, Mary Eming (ed.) (1997), *Early Child Development: Investing in Our Children's Future,* Amsterdam: Elsevier.
Young, Mary Eming (ed.) (2002), *From Early Child Development to Human Development*, Washington, DC: World Bank.

Appendix A. Estimation of treatment effects

Let $Y_i^{P_i D_i}$ be the outcome Y for child i with $P_i = 1$ if the child lives in a programme village and $P_i = 0$ otherwise; and with $D_i = 1$ if a female household member of child i participates in the programme and $D_i = 0$ otherwise. We never observe $P_i = 0$ and $D_i = 1$ at the same time. That is, people living in control villages cannot participate in the programme. To emphasize that the population in control villages does not have the opportunity to decide to participate in the programme, we will denote the outcome Y_i^{00} as $Y_i^{\bar{0}}$.

The child outcomes for participants living in a programme village, for non-participants living in a programme village, and for the control group living in control villages, Y_i^{11}, Y_i^{01} and $Y_i^{\bar{0}}$ respectively, are a function of observable characteristics X_i and unobservable characteristics U_i of the individual. The outcomes can be characterized as follows:

$$
\begin{aligned}
& Y_i^{11} = g^{11}(X_i) + U_i^{11} && \text{if } P_i = 1 \text{ and } D_i = 1 \\
& Y_i^{10} = g^{10}(X_i) + U_i^{10} && \text{if } P_i = 1 \text{ and } D_i = 0 \\
& Y_i^{\bar{0}} = g^{\bar{0}}(X_i) + U_i^{\bar{0}} && \text{if } P_i = 0
\end{aligned}
\tag{1a}
$$

We assume that the unobserved components are normally and identically distributed with mean zero. That is, we assume $E(U_i^{11} \mid X_i) = E(U_i^{10} \mid X_i) = E(U_i^{\bar{0}} \mid X_i) = 0$.

We can observe either Y_i^{11}, Y_i^{10} or $Y_i^{\bar{0}}$ for an individual but never all three at the same time. The observed outcome Y_i can be described with a switching regression model:

$$
Y_i = P_i(D_i Y_i^{11} + (1 - D_i) Y_i^{10}) + (1 - P_i) Y_i^{\bar{0}} \tag{2}
$$

Substituting the outcome functions from equation (1a) in the switching regression model (2) yields:

$$
\begin{aligned}
Y_i = {} & g^{\bar{0}}(X_i) + (g^{11}(X_i) - g^{10}(X_i)) P_i D_i + (g^{10}(X_i) - g^{\bar{0}}(X_i)) P_i \\
& + (U_i^{11} - U_i^{10}) P_i D_i + (U_i^{10} - U_i^{\bar{0}}) P_i + U_i^{\bar{0}}
\end{aligned}
\tag{3a}
$$

Using this model and omitting the i's, we can define the treatment effects as:

$$
\begin{aligned}
ATE &= E(Y^{11} - Y^{10} \mid X) && = g^{11}(X) - g^{10}(X) \\
ATET &= E(Y^{11} - Y^{10} \mid X, P = 1, D = 1) && = g^{11}(X) - g^{10}(X) + E(U^{11} - U^{10} \mid X, P = 1, D = 1)
\end{aligned}
$$

$$
\begin{aligned}
SOE &= E(Y^{10} - Y^{\bar{0}} \mid X) && = g^{10}(X) - g^{\bar{0}}(X) \\
SOET &= E(Y^{10} - Y^{\bar{0}} \mid X, P = 1, D = 0) && = g^{10}(X) - g^{\bar{0}}(X) + E(U^{10} - U^{\bar{0}} \mid X, P = 1, D = 0)
\end{aligned}
$$

$$
Total\ ATE = E(Y^{11} - Y^{\bar{0}} \mid X) = g^{11}(X) - g^{\bar{0}}(X)
$$

$$\textit{Total ATET} = E(Y^{11} - Y^{\bar{0}} \mid X, P=1, D=1) = g^{11}(X) - g^{\bar{0}}(X) + E(U^{11} - U^{\bar{0}} \mid X, P=1, D=1)$$

where *ATE* and *ATET* stand for Average Treatment Effect and Average Treatment Effect on the Treated. *SOE* and *SOET* represent the Spillover Effect and the Spillover Effect on the Treated.

To clarify the notation, we temporarily assume that –apart from treatment dummies, we have $g^{11}(X_i) = g^{10}(X_i) = g^{\bar{0}}(X_i) = g(X_i)$. In practice, introducing interaction terms between the X_i's and the P_i- and D_i-dummies will allow for different functions of the observable characteristics in the three different states. Similarly, we assume a linear functional form, to be modified introducing interaction terms and higher order terms of the X_i's. Then the outcome functions in (1a) become:

$$\text{(1b)} \qquad \begin{aligned} Y_i^{11} &= g(X_i) + \gamma + \phi + U_i^{11} && \text{if } P_i = 1 \text{ and } D_i = 1 \\ Y_i^{10} &= g(X_i) + \phi + U_i^{10} && \text{if } P_i = 1 \text{ and } D_i = 0 \\ Y_i^{\bar{0}} &= g(X_i) + U_i^{\bar{0}} && \text{if } P_i = 0 \end{aligned}$$

The switching regression model simplifies to:

$$\text{(3b)} \qquad Y_i = g(X_i) + \gamma P_i D_i + \phi P_i + (U_i^{11} - U_i^{10}) P_i D_i + (U_i^{10} - U_i^{\bar{0}}) P_i + U_i^{\bar{0}}$$

Based on the switching regression model in (3b), the econometric model to be estimated has the following form:

$$\text{(4)} \qquad Y_i = g(X_i) + \gamma P_i D_i + \phi P_i + \eta_i$$

where $g(X_i)$ captures the effect of observable characteristics on the outcome and the coefficients γ and ϕ capture shifts in the outcome due to respectively participation in the programme when living in a programme village, and to living in a programme village. In other words, since γ captures the total difference in outcomes between participants and non-participants in programme villages, ϕ represents the difference in outcomes between non-participants in programme villages and the population in control villages. The latter difference measures the spillover effect of the programme on non-participants.

In terms of the econometric model, the treatment effects are denoted as:

$$\begin{aligned} \textit{ATE} &= E(Y^{11} - Y^{10} \mid X) && = \gamma \\ \textit{ATET} &= E(Y^{11} - Y^{10} \mid X, P=1, D=1) && = \gamma + E(U^{11} - U^{10} \mid X, P=1, D=1) \\ \\ \textit{SOE} &= E(Y^{10} - Y^{\bar{0}} \mid X) && = \phi \\ \textit{SOET} &= E(Y^{10} - Y^{\bar{0}} \mid X, P=1, D=0) && = \phi + E(U^{10} - U^{\bar{0}} \mid X, P=1, D=0) \\ \\ \textit{Total ATE} &= E(Y^{11} - Y^{\bar{0}} \mid X) && = \gamma + \phi \end{aligned}$$

$$\textit{Total ATET} \quad = E(Y^{11} - Y^{\bar{0}} \mid X, P=1, D=1) \quad = \gamma + \phi + E(U^{11} - U^{\bar{0}} \mid X, P=1, D=1)$$

Which parameters can be recovered under which assumptions?

We observe the following three outcomes:

$$E(Y^{11} \mid X, P=1, D=1) = g(X) + \gamma + \phi + E(U^{11} \mid X, P=1, D=1)$$
$$E(Y^{10} \mid X, P=1, D=0) = g(X) + \phi + E(U^{10} \mid X, P=1, D=0)$$
$$E(Y^{\bar{0}} \mid X, P=0) = g(X) + E(U^{\bar{0}} \mid X, P=0)$$

A naïve comparison of the three sample groups, yields the following estimates:

1) A comparison of the observed outcomes for participants with non-participants in programme villages, i.e. $E(Y^{11} \mid X, P=1, D=1) - E(Y^{10} \mid X, P=1, D=0)$, yields the estimate $\hat{\gamma}$. Using the outcome equations (1b) and adding and subtracting an additional term, this estimate can be decomposed as follows:

   $\hat{\gamma}$ $= \gamma + E(U^{11} \mid X, P=1, D=1) - E(U^{10} \mid X, P=1, D=0)$
   $= \gamma + E(U^{11} - U^{10} \mid X, P=1, D=1) + E(U^{10} \mid X, P=1, D=1) - E(U^{10} \mid X, P=1, D=0)$
   = *ATE* + participation treatment heterogeneity within programme villages + participation selection bias
   = *ATET* + participation selection bias

   If there is participation selection bias, we need to correct for it using instrumental variables. In that case and provided that there is no participation treatment heterogeneity, $\hat{\gamma}$ will measure the average treatment effect *ATE*.

   There are two cases for which participation treatment heterogeneity is zero (Heckman, 1997). First, when there are no unobservable components of the relative gains for participants. That is, conditional on the *X*-variables, the effect of the programme is the same for everyone:
   $E(U^{11} - U^{10} \mid X, P=1, D=1) = E(U^{11} - U^{10} \mid X, P=1, D=0) = E(U^{11} - U^{10} \mid X, P=1) = 0$.
   The second case for which treatment heterogeneity equals zero, but with $U^{11} \neq U^{10}$ given *X, P* and *D*, requires that the information on individual gains does not influence the participation decision of an individual (or that individuals do not act on this information). If people do not know what will be their individual gains, their best expectation might be the population gain, which equals 0.

   If on the other hand there is both participation selection bias and participation treatment heterogeneity for participants, the instrumental variables estimate $\hat{\gamma}$ will be a biased estimate of the *ATET* (Heckman, 1997).[41] It then represents a *LATE* or Local Average Treatment Effect (Angrist et al., 1996).

[41] To see why instrumental variables will produce biased estimates in case of participation treatment heterogeneity, look at the simplified comparison of participants with non-participants (omitting the village dummy P_i). The switching regression model of (3b) becomes:
$Y_i = g(X_i) + \gamma D_i + (U_i^{\,1} - U_i^{\,0}) D_i + U_i^{\,0}$.
This can be reformulated (adding and subtracting additional components) as:

If participation selection bias is ignorable, the estimate represents the *ATE,* or the *ATET* in case of participation treatment heterogeneity.

2) A comparison of the observed outcomes for non-participants in programme villages with the population in control villages, i.e. $E(Y^{10} \mid X, P=1, D=0) - E(Y^{\bar{0}} \mid X, P=0)$, yields the estimate $\hat{\phi}$. This estimate can be decomposed as follows:

$$\begin{aligned}\hat{\phi} &= \phi + E(U^{10} \mid X, P=1, D=0) - E(U^{\bar{0}} \mid X, P=0) \\ &= \phi + E(U^{10} - U^{\bar{0}} \mid X, P=1, D=0) \\ &\quad + E(U^{\bar{0}} \mid X, P=1, D=0) - E(U^{\bar{0}} \mid X, P=1) + E(U^{\bar{0}} \mid X, P=1) - E(U^{\bar{0}} \mid X, P=0)\end{aligned}$$

= *SOE* + village treatment heterogeneity for non-participants
+ participation selection bias[42] + village selection bias
= *SOET* + participation selection bias + village selection bias

Due to the nested structure of the sample, we not only consider participation selection bias and participation treatment heterogeneity at the individual level (which is the standard approach), but also selection bias and treatment heterogeneity at the village level.

Given the discussion in section 3. we assume that village selection bias is negligible. In other words, we assume that there is no systematic difference in the effect of unobservable characteristics on child outcomes (after controlling for observables) between the population in programme and control villages if the

---

$Y_i = g(X_i) + [\gamma + E(U^1 - U^0 \mid X, D=1)]D_i + [(U_i{}^1 - U_i{}^0) - E(U^1 - U^0 \mid X, D=1)]D_i + U_i{}^0$

The comparison of participants with non-participants, $E(Y^1 \mid X, D=1) - E(Y^0 \mid X, D=0)$, yields

$\hat{\gamma} = \gamma + E(U^1 - U^0 \mid X, D=1) + E(U^0 \mid X, D=1) - E(U^0 \mid X, D=0)$

In case of selection bias, *D* is correlated with the unobserved U^0. Instrumenting for *D* can remove this bias. However, in case of treatment heterogeneity, i.e. when $E(U^1 - U^0 \mid X, D=1) \neq (U^1 - U^0)$, and thus $[U^1 - U^0 - E(U^1 - U^0 \mid X, D=1)] \neq 0$, instrumenting for *D* will remove the correlation with U^0 but not the correlation with the rest of the unobservables. Since the vector *Z* of instrumental variables is correlated with *D*, it will also be correlated with part of the error term, namely with: $[U^1 - U^0 - E(U^1 - U^0 \mid X, D=1)]$.

[42] Participation selection bias in this case is measured as the difference between the unobserved components in the control state for participants versus the average programme village inhabitant. This amounts to the same as in comparison 1) where participation selection bias was represented as the difference between the unobserved components in the control state for participants versus non-participants in programme villages. Note that:

$E(U^{\bar{0}} \mid X, P=1) = E(U^{\bar{0}} \mid X, P=1, D=0)\Pr(D=0 \mid X, P=1) + E(U^{\bar{0}} \mid X, P=1, D=1)[1 - \Pr(D=0 \mid X, P=1)]$

We have participation selection bias if $E(U^{\bar{0}} \mid X, P=1, D=0) - E(U^{\bar{0}} \mid X, P=1) \neq 0$. In other words, substituting for $E(U^{\bar{0}} \mid X, P=1)$, we have participation selection bias if

$E(U^{\bar{0}} \mid X, P=1, D=0)\,[1 - \Pr(D=0 \mid X, P=1) \neq E(U^{\bar{0}} \mid X, P=1, D=1)[1 - \Pr(D=0 \mid X, P=1)$, which means that $E(U^{\bar{0}} \mid X, P=1, D=1) - E(U^{\bar{0}} \mid X, P=1, D=0) \neq 0$.

programme had not been implemented at all. That is, $E(U^{\bar{0}} \mid X, P=1) = E(U^{\bar{0}} \mid X, P=0) = 0$. Using instrumental variables for the decision to participate will correct for potential participation selection bias.

If there is no village treatment heterogeneity for non-participants, $\hat{\phi}$ will be an unbiased estimate of the *SOE*, the spillover effect of the programme on non-participants. No village treatment heterogeneity implies that (a) the population in control villages would have benefited to the same extent as non-participants from unobserved gains from the programme if they had lived in a programme village, or (b) the programme villages are not selected on these unobserved gains for non-participants relative to control villages.

In the presence of village treatment heterogeneity, the instrumental variables estimate retrieves the *SOET* (the spillover effect on the treated), i.e. the effect of the programme on the non-participating population that actually lives in programme villages.

3) The comparison of the observed outcomes for participants in programme villages with the population in control villages, $E(Y^{11} \mid X, P=1, D=1) - E(Y^{\bar{0}} \mid X, P=0)$, yields the sum of the estimates $\hat{\gamma}$ and $\hat{\phi}$.

$$
\begin{aligned}
\hat{\gamma}+\hat{\phi} &= \gamma+\phi+E(U^{11} \mid X, P=1, D=1) - E(U^{\bar{0}} \mid X, P=0) \\
&= \gamma+\phi+E(U^{11}-U^{10} \mid X, P=1, D=1) + E(U^{10}-U^{\bar{0}} \mid X, P=1, D=1) \\
&\quad +[E(U^{\bar{0}} \mid X, P=1, D=1) - E(U^{\bar{0}} \mid X, P=1)] + [E(U^{\bar{0}} \mid X, P=1) - E(U^{\bar{0}} \mid X, P=0)]^{43} \\
&= \textit{Total ATE} + \text{participation treatment heterogeneity} \\
&\quad + \text{village treatment heterogeneity for participants} \\
&\quad + \text{participation selection bias} + \text{village selection bias} \\
&= \textit{Total ATET} + \text{participation selection bias} + \text{village selection bias}
\end{aligned}
$$

In section 3, we argue that village selection bias can be ignored. With instrumental variables and in the absence of participation and village treatment heterogeneity, we recover the total average treatment effect (*total ATE*). With instrumental

[43] The decomposition in comparison 3. can be rewritten as the sum of the decompositions in comparisons 1. and 2. After omitting the common terms $E(U^{11}-U^{10} \mid X, P=1, D=1)$, $E(U^{\bar{0}} \mid X, P=1)$ and $E(U^{\bar{0}} \mid X, P=0)$, the two versions (i.e. comparison 3. decomposition and the sum of the comparison 1. and 2. decompositions) are equal to each other if:

$$
\begin{aligned}
&E(U^{10}-U^{\bar{0}} \mid X, P=1, D=1) + E(U^{\bar{0}} \mid X, P=1, D=1) - E(U^{\bar{0}} \mid X, P=1) = \\
&E(U^{10}-U^{\bar{0}} \mid X, P=1, D=0) + E(U^{\bar{0}} \mid X, P=1, D=0) - E(U^{\bar{0}} \mid X, P=1) \\
&\quad + E(U^{10} \mid X, P=1, D=1) - E(U^{10} \mid X, P=1, D=0) .
\end{aligned}
$$

Rearranging terms, the above equality states that village treatment heterogeneity for participants equals the village treatment heterogeneity for non-participants plus additional differences in programme and control outcomes for participants versus non-participants:

$$
\begin{aligned}
&E(U^{10}-U^{\bar{0}} \mid X, P=1, D=1) = \\
&E(U^{10}-U^{\bar{0}} \mid X, P=1, D=0) + \{E(U^{10} \mid X, P=1, D=1) - E(U^{10} \mid X, P=1, D=0)\} \\
&-\{E(U^{\bar{0}} \mid X, P=1, D=1) - E(U^{\bar{0}} \mid X, P=1, D=0)\}
\end{aligned}
$$

variables and village treatment heterogeneity but no participation treatment heterogeneity, we recover the total average treatment effect on the treated. If there is participation treatment heterogeneity as well as participation selection bias, the instrumental variables estimate can be interpreted as a local average treatment effect (*LATE*).

Table 1. Description of the dependent and independent variables

Variables[44]	Description	Nr. of obs.[45]	Mean [46]	S.e. [47]	Min	Max
Dependent variables						
Preschool (only children 3-5 years of age)	Goes to preschool=1; otherwise=0	1025	.21	.034	0	1
Primary/middle school (only children 6-13 years of age)	Is currently enrolled in primary/middle school=1; otherwise=0	2516	.80	.029	0	1
Polio immunization	Immunized against polio=1; otherwise=0	4662	.88	.015	0	1
Tuberculosis immunization	Immunized against tuberculosis=1; otherwise=0	4622	.67	.027	0	1
DTP immunization	Immunized against DTP=1; otherwise=0	4594	.64	.028	0	1
Measles immunization	Immunized against measles=1; otherwise=0	4591	.52	.049	0	1
Child characteristics						
Sex	Female=1; male=0	4713	.47	.017	0	1
Age		4758	6.53	.155	0	13
Preschool (only children 6-13 years of age)	Does/did go to preschool=1; otherwise=0	2631	.16	.028	0	1
Household characteristics						
Scheduled Castes / Scheduled Tribes (SC/ST) [48]	Household belongs to SC/ST=1; otherwise=0	1991	.23	.030	0	1
Other Backward Castes (OBC)	Household belongs to OBC=1; otherwise=0	1991	.52	.026	0	1
General Castes	Household belongs to GC=1, otherwise=0	1991	.10	.020	0	1
Muslim	Household is Muslim=1; otherwise=0	1991	.12	.016	0	1
Highest education in household	Highest level for adult household members: No schooling=1 Primary school incomplete=2 Primary school complete=3 Middle school complete=4 High school complete or higher education=5	1963	3.24	.197	1	5
Highest female education in household	See highest education in household, for female household members only	1959	1.94	.108	1	5
Income quintile	Quintiles based on score from principal component analysis.[49] First quintile for poorest and fifth quintile for richest households.	1991	3.32	.143	1	5
Female head of household	Female head of household=1; male head of household=0	1991	.05	.010	0	1
Household size	Number of household members	1991	6.86	.148	1	22
Child dependency ratio	Number of children aged 0-13 divided by household size	1991	.66	.037	0	4
Land ownership	In acres	1949	3.75	.371	0	221
Community characteristics						
Village development indicator	Village development score based on principal component analysis[50]. An increasing score indicates higher village development.	102	.66	.198	0	1

[44] Child-related variables refer to all children in the age group of 0 to 13 years, unless mentioned otherwise.

[45] The total number of children aged 0 to 13 years in the sample is 4758. The total number of households in the sample is 1991, and the total number of villages is 102. The number of observations per variable may differ because of missing values or "don't know" responses.

[46] The mean of 0-1 dummies can be interpreted as a percentage.

[47] Standard errors are corrected for the survey design.

[48] The percentage of the population belonging to the Scheduled Tribes is very small in Bihar.

[49] The income score is derived from a principal component analysis based on 18 household assets and facilities: quality of walls, roof, floor, home ownership, type of fuel, sanitation facilities, water facilities, garbage disposal, electricity, landownerhip, ownership of clock, cycle, radio, television, sewing machine, motor, fridge, car. The first factor, with an eigenvalue of 4.758, is retained. The poverty index, with mean zero, is calculated using the factor loadings of the first factor. Missing values were imputed. Based on the poverty index, quintiles were constructed where the first quintile represents the poorest 20% of the households. For a motivation of using a score to proxy Indian household wealth based on weightings for household assets, see Filmer and Pritchett (2001).

Table 2. Comparison of participants and non-participants in programme villages

	Non-participants	Participants	P-value of the difference between participants and non-participants
% SC/ST	22.37	38.99	0.005***
% OBC	53.43	49.15	0.419
% General castes	10.50	2.25	0.002***
% Muslim	11.58	9.21	0.568
Average income quintile	3.37	2.89	0.002***
Average area of land owned (acres)	0.38	0.18	0.000***
Highest level of education in household	3.29	2.78	0.032**
Highest female level of education in household	1.98	1.59	0.003***
Household size	6.89	6.51	0.083*
Child dependency ratio	0.66	0.84	0.001***
Number of children <= 13 years of age	2.39	2.60	0.074*
% with female head of household	5.09	5.68	0.703

- Estimates are weighted for the stratified clustered sample design, with standard errors corrected accordingly.
- For a description of the variables, see table 1.
- *: p-value < 0.10; **: p-value < 0.05; ***: p-value < 0.01.

Table 3. Mahila Samakhya programme data by district

	Darbhanga	Muzaffarpur	Sitamarhi
Total nr. of blocks in district	18	16	17*
Total nr. of blocks where the MS programme is active	2	5	13
Total nr. of villages in MS blocks	1269	1833	846
Total nr. of villages where the MS programme is active	250	409	300
% of villages covered	19.7%	22.3%	35.5%
Nr. of active groups**	190	374	345***

Source: Mahila Samakhya District Project Offices (Darbhanga, Muzaffarpur, Sitamarhi), March 2003.

* After a recent administrative bifurcation of Sitamarhi in two districts, Sitamarhi currently has 8 blocks. Only these 8 blocks are incorporated in the sample.
** Not including emerging groups. Source: Mahila Samakhya (2002), data as of March 2002.
*** In some villages more than one group is active.

[50] The village development indicator is derived from a principal component analysis of the quality of the roads within the village, the availability within the village boundaries of a market, post office, telephone, bus, and bank, and the quality of the roads. The first factor, with an eigenvalue of 2.930, is retained and its factor loadings are used to calculate the village development indicator, imputing missing values. The score is rescaled to a minimum of 0 and a maximum of 1.

Table 4. Comparison of programme and control villages

	Control villages	Programme villages	P-value of the difference
PANEL A			
Population characteristics			
% Scheduled Castes / Scheduled Tribes	23.44	23.27	0.980
% Other Backward Castes	45.61	53.20	0.210
% General Castes	13.78	10.05	0.426
% Muslim	18.72	11.45	0.216
Average income quintile	3.13	3.34	0.266
Average land ownership (in acres)	0.35	0.37	0.786
Highest education level in household	3.07	3.26	0.473
Highest female education level in household	1.80	1.96	0.339
Average household size	6.74	6.87	0.606
Average child dependency ratio	0.60	0.67	0.191
Number of children <=13 years old	2.21	2.40	0.158
% households with female head	5.70	5.13	0.683
PANEL B			
Village characteristics			
% with flood in last three years	90.86	78.05	0.066*
% with draught in last three years	46.45	45.80	0.944
Average road quality	1.53	1.54	0.972
Availability of public transport	48.77	56.40	0.531
Average child wage (rupees)	29.01	25.03	0.152
Average distance to nearest town (km)	24.81	12.80	0.000***
% of villages with facility available			
Market	44.24	40.50	0.743
Post office	45.51	25.62	0.071*
Telephone	71.14	62.21	0.299
Bus stop	27.22	31.20	0.723
Bank	17.82	12.72	0.582
Health center	36.38	41.62	0.659
Distance from community if not available (km)			
Market	2.07	2.63	0.133
Post office	2.10	2.20	0.717
Telephone	4.57	2.85	0.313
Bus stop	2.66	2.89	0.681
Bank	3.24	3.43	0.703
Health center	3.30	2.94	0.516
PANEL C			
Average number of villages that have at least one school of the type specified:			
Preschool	24.64	66.88	0.000***
Primary school	88.06	85.50	0.760
Middle school	24.26	23.01	0.905
High school	3.34	6.44	0.526

- Estimates are weighted for the stratified clustered sample design, with standard errors corrected accordingly.
- *: p<0.10, **: p<0.05, ***: p<0.01

Table 5. Immunization rates (children aged 0 to 13 years)

PANEL A. Comparison of control villages and programme villages

		Programme village			P-values of the difference between:	
	Control village	**Programme village**	**Members only**	**Non-members only**	Control and programme village	Members and non-members
Tuberculosis	44.34	68.97	74.03	68.66	0.0001***	0.165
Polio	88.89	88.40	88.92	88.36	0.8380	0.811
DTP	41.73	65.72	70.58	65.41	0.0000***	0.174
Measles	36.24	53.94	64.50	53.26	0.0125**	0.028**

- Estimates are weighted for the stratified clustered sample design, with standard errors corrected accordingly.
- *: p<0.10, **: p<0.05, ***: p<0.01

PANEL B. Immunization by caste

	Scheduled Castes	**Other Backward Castes**	**General Castes**	**Muslim**
Tuberculosis	63.92	66.13	87.97***	45.14***
Polio	86.16	86.90	94.59**	91.41
DTP	64.83	60.95	87.98***	43.58***
Measles	57.54	51.26	77.38***	35.76***

- Estimates are weighted for the stratified clustered sample design, with standard errors corrected accordingly.
- P-values are calculated for the difference in means between a subpopulation (e.g. Scheduled Castes) and the rest of the child population. *: p<0.10, **: p<0.05, ***: p<0.01

PANEL C. Immunization by gender

	Total		**Members**		**Non-members**		**Control**	
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
Tuberculosis	66.06	67.58	75.56	72.07	67.60	69.67	45.74	42.98
Polio	89.09	87.95	90.12	87.61	88.84	88.12	90.77	86.47**
DTP	63.24	63.89	72.20	68.47	64.73	65.94	43.97	39.83
Measles	52.07	52.58	66.74	61.75	52.71	53.67	37.63	35.25

- Estimates are weighted for the stratified clustered sample design, with standard errors corrected accordingly.
- P-values are calculated for the difference in means between boys and girls in a subpopulation (e.g. members only). *: p<0.10, **: p<0.05, ***: p<0.01

Table 6. Preschool enrolment (children aged 3 to 5 years)

		Programme village			P-values of the difference between:	
	Control village	**Programme village**	**Members only**	**Non-members only**	Control and programme village	Members and non-members
PANEL A. Availability of preschools						
Average # of preschools in village	0.35	0.88			.004***	
% of villages with at least one preschool	24.64	66.88			.000***	
PANEL B. % preschool enrolment of children 3 to 5 years old						
% of all children	7.73	22.21	53.40	20.36	.008***	.000***
% of boys	8.54	18.83	51.85	17.32	.065*	.000***
% of girls	6.87	26.08	54.75	24.04	.007***	.000*
% of SC/ST	0.00	23.98	48.37	21.07	.000***	.010**
% of OBC	13.67	24.18	58.28	22.52	.180	.000***
% of General Caste	9.76	17.00	25.53	16.93	.556	.724
% of Muslims	0.00	20.69	60.00	18.16	--	.078*
% of 3 years old	6.16	18.74	44.47	17.15	.023**	.000***
% of 4 years old	5.12	21.02	61.33	18.17	.012**	.000***
% of 5 years old	11.16	26.18	54.08	24.83	.055*	.000***
PANEL C. % preschool enrolment ONLY in villages with AT LEAST ONE PRESCHOOL						
% of all children	17.86	27.16	62.15	25.11	.386	.000***
% of boys	21.57	22.87	60.35	21.14	.909	.000***
% of girls	12.57	32.42	63.24	30.14	.124	.000***
% of SC/ST	0.00	24.62	53.59	21.13	.001***	.014**
% of OBC	27.79	28.98	75.76	27.27	.926	.000***
% of General Caste	100.00	17.92	25.53	17.83	.000***	.762
% of Muslims	0.00	63.72	65.59	33.40	--	.260
% of 3 years old	14.42	25.18	52.98	23.40	.275	.001***
% of 4 years old	10.96	23.98	69.13	20.80	.250	.000***
% of 5 years old	24.98	31.38	62.75	29.95	.671	.001***

- Estimates are weighted for the stratified clustered sample design, with standard errors corrected accordingly.
- *: p<0.10, **: p<0.05, ***: p<0.01
- The difference between boys and girls is never significant (for none of the subgroups: participants, non-participants, control). This holds both for panel B. and panel C. of the table.
- The difference between the castes is never significant for participants or non-participants. Among the control group children however, General Caste children are most likely to go to preschool and SC/ST and Muslim children are least likely to go to preschool. But note that for the control group *in panel C. the 0 % for SC/ST refers to only 26 children. The 100 % for General Castes and the 0% for Muslim children refer both to a single child.*

Table 7. School enrolment (children aged 6 to 13 years)

		Programme village			P-values of the difference between:	
	Control village	**Programme village**	**Members only**	**Non-members only**	Control and programme village	Members and non-members
All children	75.54	80.19	84.95	79.87	.307	.160
Boys	82.85	80.13	87.63	79.61	.568	.035**
Girls	66.32	80.76	81.77	80.70	.020**	.821
SC/ST	55.03	76.79	77.01	76.77	.003***	.971
OBC	82.06	77.85	89.52	77.08	.522	.021**
General	90.18	98.35	88.62	98.47	.129	.082*
Muslim	73.90	80.38	93.71	79.95	.473	.132

- Estimates are weighted for the stratified clustered sample design, with standard errors corrected accordingly.
- *: p<0.10, **: p<0.05, ***: p<0.01
- The difference in school enrolment between boys and girls is highly significant in control villages but not in programme villages. However, participants are less likely to send their daughters to school than non-participants (without correcting for household characteristics).
- Scheduled Castes children are significantly less likely to be enrolled in control villages, but not in programme villages. General Castes children are significantly more likely to be enrolled than other children, in both sample groups.

Table 8. Instrumental variables

Panel A. Description of the instrumental variables (non-weighted)

	Obs.	Mean	St. dev.	Min	Max
		Civicness			
Total population	1975	.000	.735	-0.618	1.707
Participants	714	.521	.802	-0.618	1.707
Non-participants	710	-.280	.479	-0.618	1.707
Control	551	-.314	.510	-0.618	1.707
		Giving assistance			
Total population	1986	.000	.816	-2.258	0.388
Participants	716	.163	.620	-2.258	0.388
Non-participants	714	-.135	.922	-2.254	0.388
Control	556	-.036	.858	-2.254	0.388

Panel B. Comparison of the sample groups (weighted observations)

		Programme villages			P-values of the difference between:	
	Control villages	Programme villages	Members only	Non-members only	Control and programme village	Members and non-members
Civicness	-0.28	-0.25	0.77	-0.31	.687	.000***
Giving assistance	-0.06	-0.33	0.18	-0.36	.003***	.000***

- Estimates are weighted for the stratified clustered sample design, with standard errors corrected accordingly.
- *: p<0.10, **: p<0.05, ***: p<0.01

Panel C. Correlation matrix of participation and instruments (programme villages only)

	Participation in MS	Civicness	Giving assistance
Participation in MS	1.000		
Civicness	0.523	1.000	
Giving assistance	0.187	0.144	1.000

Table 9. Testing the correlation of participation with instrumental variables (probit)

Dependent variable:	Participation in Mahila Samakhya (only for programme villages)	
	Coefficient	s.e.
Instrumental variables		
Civicness	.918	.114***
Assistance	.274	.061***
Individual characteristics		
Age	-.009	.004**
Household characteristics		
SC / ST	.862	.297***
Other Backward Castes	.842	.253***
Muslim	.560	.302*
Household education	.021	.030
Female education	.001	.036
Income quintile	-.034	.030
Female household head	.217	.132
Household size	-.029	.015*
Dependency ratio	.129	.060**
Community characteristics		
Distance to town	.003	.009
Village development score	-.155	.270
1991 village literacy rate	-.004	.005
Block dummies		
District dummies		
# obs	1391	
F-test (instrumental variables)	$\chi^2(2)=69.35$	p=0.000***

- Observations are weighted for the outcome-based sampling procedure.
- Standard errors are robust and corrected for clustering at the village level.
- *: p<0.10, **: p<0.05, ***: p<0.01

Table 10. Testing the instrumental variables in control villages (probit)

	Tuberculosis		DTP		Measles		Preschool (only villages with >=1 preschool)		School	
	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.
Instrumental variables										
Civicness	.198	.180	.102	.150	.195	.137	-8.307	4.658*	.240	.137*
Assistance	-.021	.103	-.087	.106	-.127	.110	-1.074	.914	.082	.063
Child characteristics										
Sex	-.118	.096	-.142	.100	-.086	.100	2.147	1.619	-.547	.142***
Age	-.058	.016***	-.057	.015***	-.041	.016***	1.061	.690	.270	.025***
Prior enrolment in preschool									.369	.288
Household characteristics										
SC/ST	-.339	.391	-.540	.379	-.693	.366*			-.247	.319
OBC	.142	.352	-.050	.354	.027	.305			.082	.210
Muslim	-.436	.403	-.338	.432	-.475	.388			-.135	.262
Household education	.086	.058	.065	.054	.107	.051**	6.553	4.195	.088	.058
Female education	.087	.077	.088	.079	.117	.077*	-3.233	2.124	.072	.046
Income quintile	.152	.054***	.173	.062***	.097	.068	.817	.563	.150	.054***
Female household head	-.045	.234	.001	.227	-.384	.281	…		.109	.351
Household size	-.045	.030	-.054	.035	-.040	.036	-.244	.199	.005	.025
Dependency ratio	-.190	.099*	-.286	.099***	-.103	.077	12.883	7.960	.099	.105
Community characteristics										
Distance to health center	-.045	.074	-.015	.079	-.112	.083				
Village/block/district variables										
Observations	1102		1097		1099		52		848	
Test of joint significance of instrumental variables	$\chi^2(2)$= 1.37	p=0.504	$\chi^2(2)$= 1.31	p=0.518	$\chi^2(2)$= 3.67	p=.160	$\chi^2(2)$= 3.21	p=0.201	$\chi^2(2)$= 4.15	p=.126

*: p<0.10, **: p<0.05, ***: p<0.01. Robust standard errors corrected for clustering at the village level.

Table 11. Summary of instrumental variables

	Tuberculosis	DTP	Measles	Preschool	School
Civicness	X	X	X	X	X
Assistance		X	X	X	X
Test of overidentifying restrictions (using 2SIV) if more than one instrument		$\chi^2(1) = 1.340$ p=0.247	$\chi^2(1) = .785$ P=0.367	$\chi^2(1) = .284$ P=.594	$\chi^2(1) = 3.229$ P=.072

Where X means that the row- variable will be used as instrument for the estimation of the particular child outcome.

Table 12. Probit estimation of immunization (without instrumental variables)

Dependent variable:	Tuberculosis		DTP		Measles		Polio	
	Coefficient	s.e.	Coefficient	s.e.	Coefficient	s.e.	Coefficient	s.e.
Programme variables								
Participation variable	.258	.091***	.304	.088***	.290	.086***	.015	.095
Programme village	.627	.279**	.547	.216**	.922	.266***	-.014	.202
Child characteristics								
Sex	-.154	.043***	-.158	.042***	-.142	.040***	-.171	.057***
Age	-.050	.008***	-.042	.008***	-.024	.008***	-.181	.014***
Household characteristics								
SC/ST	-.248	.199	-.464	.191**	-.291	.184	-.275	.175
OBC	-.170	.187	-.320	.180*	-.127	.160	-.288	.175
Muslim	-.585	.198***	-.592	.198***	-.469	.182**	-.391	.170**
Household education	.066	.024***	.065	.025***	.086	.026***	-.020	.027
Female education	.152	.034***	.157	.038***	.131	.033***	.045	.035
Income quintile	.096	.029***	.089	.031***	.076	.031**	.119	.032***
Female household head	-.165	.140	-.143	.141	-.120	.145	-.361	.158**
Household size	-.044	.013***	-.040	.014***	-.032	.014**	-.006	.015
Dependency ratio	-.135	.050***	-.166	.056***	-.057	.051	-.010	.058
Community characteristics								
Distance to town	.001	.006	.003	.006	.003	.006	.013	.005**
Distance to health center	-.021	.022	-.036	.022	-.033	.022	-.011	.021
Village development	-.092	.279	.031	.274	.246	.301	.257	.216
1991 village literacy	.002	.006	.002	.005	-.006	.006	.007	.005
Sitamarhi district	.471	.192**	.540	.203***	.217	.186	.758	.122***
Darbhanga district	-.738	.412*	-1.032	.383***	-.784	.468*	.136	.167
Block dummies								
# obs	4517		4489		4486		4556	

*: p<0.10, **: p<0.05, ***: p<0.01. Robust standard errors corrected for clustering at the village level.

Table 13. Tuberculosis with different estimation methods

Dependent variable: **Tuberculosis**	Probit (no IV) (1)		Probit (no IV; marginal prob.) (2)		Lin prob (no IV) (3)		2S IV (4)		2S IV – GIRLS only (5)		2S IV – BOYS only (6)		2S IV – Scheduled Castes (7)	
	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.
Programme variables														
Programme member	.258	.091***	.099	.034***	.084	.027***	.205	.027***	.232	.044***	.189	.034***	.118	.060*
Programme village	.627	.279**	.245	.107**	.204	.083**	.149	.049***	.112	.072**	.184	.069***	.359	.278
Child characteristics														
Sex	-.154	.043***	-.060	.017***	-.046	.013***	-.052	.012***					-.057	.022**
Age	-.050	.008***	-.020	.003***	-.015	.002***	-.015	.002***	-.013	.002***	-.017	.003***	-.013	.003***
Household char.														
SC/ST	-.248	.199	-.097	.078	-.071	.055	-.094	.027***	-.101	.041**	-.093	.035**		
OBC	-.170	.187	-.066	.072	-.042	.050	-.058	.025**	-.064	.039	-.054	.032*		
Muslim	-.585	.198***	-.230	.076***	-.181	.058***	-.176	.027***	-.174	.044***	-.182	.036***		
Household education	.066	.024***	.026	.009***	.022	.008***	.021	.005***	.017	.008**	.023	.007***	-.007	.009
Female education	.152	.034***	.059	.013***	.040	.009***	.042	.005***	.048	.007***	.035	.008***	.048	.013***
Income quintile	.096	.029***	.037	.011***	.030	.009***	.033	.006***	.041	.008***	.025	.008***	.027	.011**
Female hh head	-.165	.140	-.065	.056	-.043	.044	-.046	.032	-.042	.045	-.049	.046	-.269	.067***
Household size	-.044	.013***	-.017	.005***	-.013	.004***	-.010	.003***	-.006	.004	-.015	.004***	.009	.006
Dependency ratio	-.135	.050***	-.052	.019***	-.039	.016**	-.044	.010***	-.041	.015***	-.049	.014**	-.079	.019***
Community char.														
Distance to town	.001	.006	.000	.002	-.000	.002	-.000	.001	-.003	.001**	.002	.002	.003	.002
Distance health center	-.021	.022	-.008	.008	-.006	.006	-.005	.004	-.003	.005	-.007	.005	-.043	.010***
Village development	-.092	.279	-.036	.109	-.021	.087	-.039	.046	.132	.071*	.047	.059	-.285	.090***
1991 village literacy	.002	.006	.001	.002	.001	.002	.001	.001	.001	.001	.001	.001	-.003	.001***
Sitamarhi district	.471	.192**	.179	.071**	.146	.061**	.137	.085	.142	.127	.074	.155	.183	.256
Darbhanga district	-.738	.412*	-.286	.155*	-.240	.125*	-.466	.129**	-.251	.190	-.476	.149***	-.166	.335
Block dummies														
# obs	4517		4517		4517		4469		2117		2352		1256	
R-squared					.271		.260		.285		.241		.296	

*: p<0.10, **: p<0.05, ***: p<0.01. Robust standard errors, corrected for clustering at the village level, and bootstrapped with 200 replications in the 2S IV estimations.

Table 14. DTP with different estimation methods

Dependent variable:	Probit (no IV)		Lin prob (no IV)		2S IV		2S IV – GIRLS only		2S IV – BOYS only		2S IV – Scheduled Castes	
DTP	(1)		(3)		(4)		(5)		(6)		(7)	
	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.
Programme variables												
Participation	.304	.088***	.097	.026***	.209	.028***	.252	.044***	.176	.034***	.135	.058**
Programme village	.547	.216**	.172	.067**	.122	.050**	.105	.073**	.135	.070*	-.108	.352
Child characteristics												
Sex	-.158	.042***	-.047	.012***	-.052	.012***					-.036	.023
Age	-.042	.008***	-.013	.002***	-.012	.002***	-.011	.003***	-.014	.003***	-.012	.003***
Household char.												
SC/ST	-.464	.191**	-.125	.051**	-.146	.028***	-.146	.043***	-.155	.035***		
OBC	-.320	.180*	-.077	.047	-.093	.025***	-.103	.040**	-.085	.032**		
Muslim	-.592	.198***	-.170	.056***	-.166	.028***	-.170	.044***	-.169	.037***		
Household education	.065	.025***	.022	.008***	.020	.005***	.017	.008**	.022	.007***	-.002	.009
Female education	.157	.038***	.045	.011***	.047	.006***	.059	.008***	.035	.008***	.023	.016
Income quintile	.089	.031***	.026	.009***	.029	.006***	.029	.008***	.028	.008***	.049	.011***
Female hh head	-.143	.141	-.040	.043	-.043	.031	-.049	.041	-.038	.045	-.226	.069***
Household size	-.040	.014***	-.011	.004***	-.009	.003***	-.003	.004	-.015	.004***	.004	.005
Dependency ratio	-.166	.056***	-.045	.016***	-.052	.009***	-.038	.014***	-.068	.014***	-.072	.017***
Community char.												
Distance to town	.003	.006	.001	.002	.000	.001	-.002	.002	.003	.002*	.008	.002**
Distance health center	-.036	.022	-.010	.007	-.010	.004***	-.014	.005**	-.006	.005	-.046	.010***
Village development	.031	.274	.020	.085	.001	.043	-.097	.067	.081	.058	-.158	.088*
1991 village literacy	.002	.005	.001	.002	.001	.001	.001	.001	.001	.001	-.004	.001***
Sitamarhi district	.540	.203***	.136	.066**	.158	.096*	.321	.152**	.097	.161	-.492	.414
Darbhanga district	-1.032	.383***	-.517	.073***	-.311	.106***	-.305	.179*	-.310	.123***	-.458	.312
Block dummies												
# obs	4489		4489		4435		2102		2333		1238	
R-squared			.294		.283		.291		.281		.335	

*: p<0.10, **: p<0.05, ***: p<0.01. Robust standard errors, corrected for clustering at the village level, and bootstrapped with 200 replications in the 2S IV estimations.

Table 15. Measles with different estimation methods

Dependent variable:	Probit (no IV)		Lin prob (no IV)		2S IV		2S IV – GIRLS only		2S IV – BOYS only		2S IV – Scheduled Castes	
Measles	(1)		(3)		(4)		(5)		(6)		(7)	
	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.
Programme variables												
Participation	.290	.086***	.092	.027***	.185	.027***	.199	.046***	.170	.039***	.031	.058
Programme village	.922	.266***	.309	.084***	.268	.050***	.273	.068***	.267	.070***	.157	.339
Child characteristics												
Sex	-.142	.040***	-.043	.012***	-.045	.013***					-.019	.024
Age	-.024	.008***	-.008	.002***	-.007	.002***	-.009	.002***	-.006	.003**	-.007	.003**
Household char.												
SC/ST	-.291	.184	-.086	.054	-.106	.028***	-.072	.044	-.141	.036***		
OBC	-.127	.160	-.034	.047	-.046	.026*	-.016	.042	-.069	.032**		
Muslim	-.469	.182**	-.141	.054**	-.136	.030***	-.120	.045**	.156	.039***		
Household education	.086	.026***	.027	.008***	.026	.005***	.030	.008***	.023	.007***	.004	.008
Female education	.131	.033***	.043	.010***	.045	.006***	.054	.009***	.036	.009***	.022	.017
Income quintile	.076	.031**	.024	.009**	.025	.006***	.016	.008**	.031	.008***	.027	.010**
Female hh head	-.120	.145	-.035	.044	-.037	.029	-.012	.037	-.050	.044	-.233	.057***
Household size	-.032	.014**	-.009	.004**	-.008	.003***	-.003	.004	-.012	.003***	.011	.005**
Dependency ratio	-.057	.051	-.013	.015	-.019	.009**	-.006	.014	-.030	.014*	-.024	.017
Community char.												
Distance to town	.003	.006	.001	.002	.000	.001	-.001	.001	.002	.002	.005	.002**
Distance health center	-.033	.022	-.010	.007	-.009	.004**	-.010	.005**	-.007	.005	-.046	.010***
Village development	.246	.301	.073	.092	.044	.048	.012	.067	.086	.061	-.113	.087
1991 village literacy	-.006	.006	-.001	.002	-.001	.001	-.001	.001	-.002	.001**	-.004	.001***
Sitamarhi district	.217	.186	.072	.059	.174	.085**	.055	.162	.065	.127	-.312	.426
Darbhanga district	-.784	.468*	-.534	.068***	-.198	.146	-.194	.230	-.208	.144	-.360	.294
Block dummies												
# obs	4486		4486		4432		2100		2332		1239	
R-squared			.291		.283		.275		.270		.361	

*: p<0.10, **: p<0.05, ***: p<0.01. Robust standard errors, corrected for clustering at the village level, and bootstrapped with 200 replications in the 2S IV estimations.

Table 16 Preschool enrolment with different estimation methods

	ALL VILLAGES		ONLY VILLAGES WITH AT LEAST ONE PRESCHOOL									
Dependent variable:	Lin. Prob. (no IV)		Probit (no IV)		Lin prob (no IV)		2S IV		2S IV (Girls only)		2S IV (Boys only)	
Preschool enrolment	(1)		(3)		(4)		(5)		(6)		(7)	
	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.
Programme variables												
Participation	.174	.036***	.703	.159***	.213	.051***	.214	.071***	.146	.107	.215	.113*
Programme village	.204	.057***	1.451	.379***	.444	.158***	.443	.075***	.572	.092***	.368	.110***
Child characteristics												
Sex	.024	.027	.111	.124	.036	.039	.023	.041				
Age	.027	.014*	.114	.073	.037	.023	.041	.023	.053	.033	.037	.032
Household char.												
SC/ST	-.018	.064	-.445	.421	-.133	.143	-.129	.118	.050	.204	-.301	.155
OBC	.052	.056	.022	.410	-.004	.141	-.012	.116	-.160	.192	-.162	.152
Muslim	-.007	.069	-.621	.550	-.169	.183	-.174	.130	-.029	.211	-.348	.174
Household education	.011	.011	.032	.054	.010	.017	.010	.016	-.007	.027	-.024	.020
Female education	.001	.013	-.073	.056	-.024	.018	-.020	.018	.001	.030	-.047	.031
Income quintile	-.002	.012	.039	.069	-.008	.021	-.006	.017	-.025	.026	.002	.026
Female head of h.h.	-.089	.050*	-.266	.320	-.070	.104	-.070	.111	.096	.136		
Household size	.001	.006	.033	.029	.008	.009	.007	.008	.006	.014	.002	.011
Dependency ratio	.019	.025	.201	.116*	.063	.034*	.054	.029*	.085	.047*	.017	.041
Community char.												
Distance to town	-.000	.003	-.008	.015	-.002	.005	-.003	.003	.003	.006	-.008	.004**
Village development	.096	.129	-.448	.656	-.147	.204	-.226	.120*	-.066	.023**	-.351	.157**
# preschools	.178	.040***	.478	.350	.134	.117	.169	.060**	.196	.084	.115	.085
1991 village literacy	.002	.002	-.000	.010	-.000	.003	.001	.002	.001	.004	.000	.003
Sitamarhi district	.139	.110	1.670	.572***	.495	.180***	.537	.111***	.496	.154***	.571	.153***
Darbhanga district	-.249	.099**	-.563	.256**	-.205	.152	-.195	.088**	-.149	.139	-.224	.120*
Block variables												
# obs	1002		510		518		509		247		262	
R-squared	.272				.298		.307		.319		.344	

*: p<0.10, **: p<0.05, ***: p<0.01. Robust standard errors, corrected for clustering at the village level, and bootstrapped with 200 replications in the 2S IV estimations.

Table 17. School enrolment with different estimations methods

Dependent variable: **School enrolment**	Probit (no IV) (1)		Probit (no IV) (1)		Lin prob (no IV) (2)		2S IV (3)		2S IV – GIRLS (4)		2S IV – BOYS (5)		2S IV – SC/ST (6)	
			Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.
Programme variables														
Participation	.298	.086***	.304	.085***	.072	.023***	.145	.036***	.111	.060*	.181	.042***	.091	.081
Programme village	-.059	.124	.324	.204	.081	.056	.046	.034	.170	.061***	-.051	.045	.202	.099**
Int. village & literacy			-.030	.012**	-.007	.003**	-.007	.002***	-.010	.003***	-.005	.003*	-.013	.004***
Child characteristics														
Sex	-.339	.087***	-.338	.088***	-.085	.023***	-.084	.017***					-.044	.034
Age	.069	.015***	.069	.015***	.017	.004***	.018	.004***	.022	.005***	.015	.005***	.026	.009***
Preschool	.177	.109	.169	.111	.055	.027**	.050	.019***	.032	.029	.056	.027**	.032	.040
Household characteristics														
SC/ST	-.480	.187**	-.451	.194**	-.089	.045**	-.100	.034***	-.083	.052	-.124	.041***		
OBC	-.299	.163*	-.299	.170*	-.045	.032	-.053	.029*	-.080	.045*	-.032	.034		
Muslim	-.366	.181**	-.351	.188*	-.064	.045	-.062	.034*	-.046	.053	.075	.043*		
Household education	.135	.030***	.133	.030***	.036	.008***	.035	.006***	.041	.009***	.031	.008***	.028	.012**
Female education	.096	.047**	.098	.048**	.008	.007	.007	.006	.022	.009**	-.006	.008	-.005	.022
Income quintile	.158	.036***	.156	.036***	.041	.010***	.042	.008***	.037	.011***	.043	.010***	.056	.014***
Female hh head	-.149	.199	-.154	.202	-.030	.057	-.034	.038	-.060	.058	.000	.050	-.151	.091
Household size	-.036	.017**	-.035	.017**	-.007	.004*	-.006	.003**	-.008	.005	-.005	.004	-.014	.008*
Dependency ratio	.012	.060	.007	.059	-.001	.017	-.005	.014	.008	.020	-.014	.017	.032	.028
Community characteristics														
Distance to town	.004	.009	.004	.009	.001	.002	.000	.001	.001	.002	.001	.001	.002	.003
# primary schools	-.106	.068	-.081	.070	-.026	.020	-.024	.013*	-.028	.020	-.024	.017	-.019	.021
Child wage	-.004	.007	-.001	.007	.000	.002	-.000	.001	.001	.002	-.001	.002	-.001	.003
Village development	-.176	.315	-.314	.319	-.065	.080	-.071	.044	-.099	.069	-.055	.059	-.225	.132*
1991 village literacy	.002	.005	.024	.011**	.006	.003**	.006	.003**	.009	.005**	.004	.002	.010	.006*
Block/district variables														
# obs	2447		2447		2447		2423		1104		1319		655	
R-squared					.138		.132		.148		.123		.131	

*: p<0.10, **: p<0.05, ***: p<0.01. Robust standard errors, corrected for clustering at the village level, and bootstrapped with 200 replications in the 2S IV estimations.

Table 18. Overview of estimated treatment effects, including Propensity Score Matching results

		Probit (marginal effect)	Linear prob. (no IV)	2SLS with IV	Total village estimation (lin. prob.)	PSM (part. vs. non-part.)	PSM (part. / non-part. vs. matched control)
		(i)	(ii)	(iii)	(iv)	(v)	(vi)
IMMUNIZATION							
Tuberculosis	Participation	.076 (.025)***	.084 (.027)***	.205 (.027)***		.088 (.026)***	.202 (.026)***
	Village	.193 (.054)***	.204 (.083)**	.149 (.049)***	.195 (.113)*		.108 (.025)***
Polio	Participation	*.004 (.000)****	*.007 (.015)*			.009 (.017)	.018 (.019)
	Village	*-.013 (.001)****	*-.008 (.032)*		-.008 (.041)		.004 (.018)
DTP	Participation	.089 (.029)***	.097 (.026)***	.209 (.028)***		.100 (.023)***	.203 (.028)***
	Village	.164 (.053)***	.172 (.067)**	.122 (.050)**	.160 (.100)		.090 (.023)***
Measles	Participation	.085 (.026)***	.092 (.027)***	.185 (.027)***		.085 (.025)***	.220 (.026)***
	Village	.262 (.100)***	.309 (.084)***	.268 (.050)***	.296 (.112)**		.089 (.024)***
PRESCHOOL							
(only villages with at least one preschool))	Participation	.212 (.053)***	.213 (.051)***	.214 (.071)***		.164 (.061)***	.340 (.080)***
	Village	.344 (.167)**	.444 (.158)***	.443 (.075)***	.382*** (.068)		.185 (.067)***
SCHOOL[51]							
School (total)	Participation	.069 (.035)**	.072 (.023)***	.145 (.036)***		.077 (.027)***	.114 (.031)***
	Village	.089 (.035***	.081 (.056)	.046 (.034)	.124 (.050)**		.056 (.031)*
School (girls)	Participation	.062 (.030)**	.063 (.029)***	.111 (.060)*		.099 (.038)***	.144 (.043)***
	Village	.185 (.064)**	.238 (.096)***	.170 (.061)***	.178 (.072)**		.139 (.042)***
School (boys)	Participation	.078 (.042)*	.082 (.029)***	.181 (.042)***		.067 (.033)**	.202 (.042)***
	Village	.001 (.000)	-.020 (.058)	-.051 (.045)	.083 (.062)		.051 (.037)

- Robust standard errors corrected for clustering at the village level. Results for IV and PSM are bootstrapped with 200 replications.
- *: p<0.10, **: p<0.05, ***: p<0.01

[51] All regression estimates for impact on school enrolment include an interaction term between programme village and 1991 village female literacy rates.