Socio-Economic Effects of a Self-Help Group Intervention:

Evidence from Bihar, India

Upamanyu Datta, the World Bank

Abstract

Poverty reduction via formation of community based organizations is a popular approach in regions of high socio-economic marginalization, especially in South Asia. The shortage of evidence on the impacts of such an approach is an outcome of the complexity of these projects, which almost always have a multi-sectoral design to achieve a comprehensive basket of aims. In the current research, we consider results from a rural livelihoods program in Bihar, one of India's poorest states. Adopting a model prevalent in several Indian states, the Bihar Rural Livelihoods Project, known locally as JEEViKA, relies on mobilizing women from impoverished, socially marginalized households into Self Help Groups. Simultaneously, activities such as micro-finance and technical assistance for agricultural livelihoods are taken up by the project and routed to the beneficiaries via these institutions; these institutions also serve as a platform for women to come together and discuss a multitude of the socio-economic problems that they face. We use a retrospective survey instrument, coupled with PSM techniques to find that JEEViKA, has engendered some significant results in restructuring the debt portfolio of these households; additionally, JEEViKA has been instrumental in providing women with higher levels of empowerment, as measured by various dimensions.

1. Introduction

It is well recognized that poverty may be caused by external shocks, but are perpetuated by unavailability of credit, malnutrition, inadequate coverage against future shocks and limited access to stable sources of income, among other factors. Such factors contribute to a self-reinforcing vicious cycle of poverty, and it is obvious that policy makers would realize that to break this cycle, a multi-sectoral approach is necessary.

It is worth noting that having the expertise to tackle each factor may be beyond a particular project. This implies that a possible multi-sectoral design must involve several entities, build synergies among them, and have a high-powered top management guiding this 'development consortium'. The other approach is to identify a 'nodal' entity which has core competencies in some of the key interventions, and ensure the liaison of other entities with the first to converge on other interventions. It is not necessary that the other entities be NGOs; one can imagine a situation that these are institutional platforms of the poor created by the 'nodal' entity to articulate demands for poverty reduction. The maintained hypothesis is that these institutional platforms will identify the key stumbling blocks to socio-economic improvement and would demand appropriate remedies from the nodal entity.

International donors and governments have realized that the 2nd approach lends itself to more sustainable project designs and have invested billions of dollars in creating such 'nodal' entities, designing subsequent interventions and finally routing benefits to last-mile beneficiaries via their institutional platforms. Indeed, various states in India have such projects functional from the last decade, which in turn led to the establishment of the country-wide National Rural Livelihoods Mission (NRLM) in 2011. In 10 years, NRLM proposes to reach out to 600000 villages of India.

Designing rigorous evaluations to understand the effects of such large scale, complex and non-standard interventions is a complicated process in itself. For example, how does one define "treatment units", when the definition of treatment itself varies across communities? Or how does one identify the appropriate "control units', given that apparent control areas are subject to substantial spillover effects, for example, in self mobilization into institutional platforms or adoption of non-financial knowledge products from treatment areas?

Perhaps, this is the main reason for the disproportionate paucity of evidence on the effects of these projects, given the variety of such projects that are currently operational. The completed researches till date usually are restricted to have a non-gold standard design, and the evidence from such studies is decidedly mixed (Mansuri & Rao, 2012). Park & Wang found no impact on the mean consumption and income of poor households but found higher consumption and income for rich households in China's Poor Village Investment Programme (Park & Wang, 2010). An evaluation of the Kecamatan Development Programme in Indonesia found positive impacts on consumption incomes for households near the poverty line, but not for more poor or disadvantaged households (Voss, 2008). Southwest China Poverty Reduction Programme led to sustained income gains only for those households that were initially poor but were relatively well educated; while the income gains for other (poor, but less educated) households faded after the lifetime of the project (Chen, Mu, & Ravallion, 2008). In the context of South Asia, the evaluation of Andhra Pradesh District Poverty Initiatives Project (APDPIP) evaluation finds positive impact on consumption and nutritional intake limited only for Self-help Group (SHG) members (Deininger & Liu, 2009).

A large literature, both theoretical and empirical, in development microeconomics, suggests that credit constraints limit income and consumption growth and increase vulnerability among poor households; when credit is routed through women, the household as a whole experiences better outcomes in the form of increased consumption or investment on goods with a public flavor. Pitt and Khandker (1998) examine 3 group based credit programs by BRAC, BRDB and GRAMEEN and find that credit routed through women increases labor supply across gender, schooling across gender, consumption expenses by the household and non-land assets held by women. Bobonis (2009) finds a similar effect of increased income for women (due to the PROGRESSA program) on expenditure for children's goods. However, Banerjee et al (2010) do not find any effects on long term investments (health, education and empowerment) due to the SPANDANA program in the urban slums of Hyderabad in Andhra Pradesh, India. Feigenberg et al (2010) find evidence in West Bengal, India that increased interaction in a group setting (for the purpose of microfinance) enhance social networking and cooperative outcomes like regular repayments and repeated credit dosage.

However, it is unclear if such programs affect women's empowerment. The complexity of measuring women's empowerment is probably a major reason why there is no clear answer. Kabeer (1999) and Agarwal (1997) provide excellent discussions about how multiple dimensions like agency, ability to choose and participation in decision making indicate women's empowerment; the authors also discuss initiatives which could affect some or all of these dimensions.

In the current research, we consider a multi-sectoral approach which closely resembles the APDPIP design. We take a close look at the impacts of a rural poverty reduction program in Bihar, one of India's poorest states. This program JEEViKA, focusses on building Self Help Groups (SHGs) of marginalized women; these groups are then federated into higher order institutions of such women at the village and local level. Cheap credit for a variety of purposes, technical assistance for various livelihood activities and encouraging awareness about various public services are the key agendas of this program. However, due to the very nature of JEEViKA's target population, and given Bihar's vicious income and gender inequality, the potential for impacts on women's empowerment exists. A retrospective survey instrument, coupled with 'Propensity Score Matching' methods are used to estimate the impacts.

The results from the survey point out that JEEViKA has played an instrumental role in restructuring the debt portfolio of beneficiary households; households that have SHG members have a significantly lower high cost debt burden, are able to access smaller loans repeatedly and borrow more often for productive purposes, when compared to households without SHG members. Since JEEViKA works by mobilizing marginalized women into institutional platforms, such women demonstrate higher levels of empowerment, when empowerment is measured by mobility, decision making and collective action. Finally, we see some effects on the asset positions, food security and sanitation preferences of beneficiary households. It is worth pointing out here that the extent and significance of the results on debt portfolio and empowerment are robust to various matching modules and various specifications of the matched sample. The results on the other dimensions are subject to specifications or matching modules.

This brings out to the point about the timeline of these interventions and the materialization of impacts. In the context of such iterative, multi-sectoral poverty reduction approach, a well-designed research question must be able to identify the goals that a project should have achieved,

given the time-line of that evaluation; the extent of such achievements are only a part of the evaluation agenda. The short review provided above provides some clues that a regular evaluation horizon of 2/3 years may be insufficient time to observe higher order effects, especially since actual benefits happen only after poor are mobilized into institutions and institutions are federated into higher-order institutions; indeed, the village-level institution, the Village Organization, which is made of 15 SHGs on an average, becomes functional 8-10 months after JEEViKA enters a village for the first time. The retrospective nature of the survey instrument also rules out any meaningful comparison of consumption or income levels between treatment and control areas.

In the view of such restrictions, it is useful to point out that this current research may be viewed as a pilot of a much more comprehensive 'multi-disciplinary' evaluation design which is now underway at JEEViKA. Thus, following the completion of this survey in early 2011, a baseline survey was conducted in 180 panchayats, located in 17 blocks of 6 districts of Bihar in mid to late 2011. After the analysis of the baseline data, JEEViKA rolled out randomly to 90 'treatment' panchayats. Allied to the design of the Randomized Control Trial, an in-depth qualitative study of 12 villages (part of the 180 panchayats) was also commissioned to look at the intervention timeline and the process of change in the villages. Finally, a behavioral study is also underway to tease out the intra and inter household effects of creating a platform to raise demand, among households and women who have otherwise faced vicious marginalization. This basket of evaluation designs is a direct outcome of the current research, which pointed out the severe restrictions that a solely quantitative approach has in understanding projects of such complexity.

In Section 2, we look at the program in greater details, including its geographical coverage, focus areas for rural development and expansion strategies. In Section 3, we discuss the design of the current research study including the most important process of identifying good counterfactual villages for the project villages, the survey instrument and the key algorithms used for propensity score matching. We consider the quality of the matched sample and discuss how different specifications of the outcome variables could give us precise estimates of the final outcome. In Section 4, we discuss the entire basket of changes that have been brought on by JEEViKA in the 6 project districts of rural Bihar. We conclude by summarizing the results and discuss future scopes of research in Section 5.

2. An Introduction to JEEViKA

Historically, Bihar has been one of India's most impoverished states, languishing at the bottom of the heap along various socio-economic dimensions. Social segregation along caste lines, gender discrimination, poor infrastructure and a near breakdown in provision of public amenities had accentuated the abysmal income levels, especially in rural Bihar. However, in recent times, Bihar has witnessed a steady turnaround under a slew of administrative reforms. In late 2006, the Govt. of Bihar inaugurated the Bihar Rural Livelihoods Project or JEEViKA, executed by the autonomous Bihar Rural Livelihoods Promotion Society and funded by the World Bank. JEEViKA slowly became the flagship rural poverty reduction program of the government, operating in 9 out of 34 districts of Bihar. Recently, JEEViKA received the mandate of scaling up its model across Bihar under the National Rural Livelihoods Mission (NRLM). Over a period of the next 10 years, the mandate is to mobilize 12.5 million rural HHs into 1 million SHGs (Self Help Group), 65000 VOs (Village Organization) and 1600 CLFs (Cluster Level Federation).

The project has certain key features, which include

a) Focusing on the poor and vulnerable members of the community, particularly women.
b) Building and empowering pro-poor institutions and organizations.
c) Emphasis on stimulating productivity growth in key livelihood sectors and employment generation in the project area.
d) Positioning project investments to be catalytic in nature to spur public and private investment in the livelihood areas/sector of poor households.
e) Identification of existing innovations in various areas and help in developing processes, systems and institutions for scaling up of these innovations.

The basic building block of the project is to promote socio-economic inclusion of rural impoverished households by mobilizing women members from such families into SHGs (Self Help Groups). In Bihar, the sharp caste segregation implies a considerable correlation between belonging to a low caste and being impoverished; additionally, in an average village in rural Bihar, low caste populations live in a separate hamlet (which may be a fair distance from the actual village center) inside the village. JEEViKA does not conduct any baseline of any kind to identify its target population; project personnel take advantage of the geographical and economic

segregation to approach the relevant hamlets and target low caste households for initial mobilization.

In an average SHG, members meet regularly to participate in savings, borrowing and repayments; additionally, it provides a small platform for 10-15 women of similar backgrounds to come together and discuss their day-to-day lives. The microfinance activities have a humble beginning where each member makes a weekly saving to the tune of 10-20 cents; the members start inter-loaning among one another, by drawing on the aggregate savings parked at the SHG. Once such practices continue over time, the project provides the SHG with a one-time grant of 900 USD, which the SHG disburses as loans to the members. Going forward, these SHGs get linked to banks and leverage funds from formal credit institutions. All avenues of such micro credit have an annual cost of 24%, as opposed to the credit from village money lenders and shopkeepers which are usually to the tune of 60% or 120% annually.

Once a minimum number of SHGs form in a village, they are federated into a Village Organization (VO); a VO is perhaps the key institution of the project as it is large enough to affect changes in the village and small enough to account for the demands coming out of the community. Thus, the key interventions of the project, such as food security, health and nutrition, livelihood activities, identification and training of youth and convergence with other schemes are driven by the VO. The VO also has a mandate to identify issues at the village level and liaison with the project's staff to provide practical solutions.

JEEViKA piloted initially in 5 blocks (sub-districts) and had its first major expansion in 2008, when it rolled out in 13 more blocks; thus at various points of times in 2008, JEEViKA started operations in 18 blocks across 6 districts of Bihar, namely, Gaya, Khagaria, Madhubani, Muzaffarpur, Nalanda and Purnea. The objective of the following study was to understand the changes brought about by the project in the socio-economic conditions of beneficiaries over a time period of 3 years, from early 2008 to end 2010.

Given JEEViKA's thrust on building institutions and providing cheap credit, we should expect that the program have impacts on debt reduction; if financial wisdom (encouraged by the program) is practiced by beneficiary households, we hope to see some movement towards credit for productive purposes. To encourage livelihood opportunities, JEEViKA's main thrust was to

provide technical assistance for agriculture; thus, we could expect to see some increased adoption of agricultural activities. Indeed, if such adoptions are significant, we may expect to see increased land holding or land leasing. Finally, given that JEEViKA beneficiaries meet weekly to engage in financial transactions and discuss agendas about their personal and communal life, we could expect that some effects on women's empowerment should be visible.

The main complication that the research team and the project team faced was that no baseline instrument was fielded prior to the expansion. Additionally, the project did not expand into the new blocks in a haphazard way; rather, the project targeted villages for entry that had large numbers of target populations. Thus, non-availability of information at baseline combined with non-random expansion complicated any interpretation of causality.

To address the problem of non-availability of data at baseline, a questionnaire with current and retrospective modules was administered in early 2011, which probed for situations at the end of 2010 and at the end of 2007. The non-random nature of JEEViKA's expansion was taken advantage of, by selecting villages from un-entered blocks (in the same districts as the entered 18 blocks) which would have been entered (according to JEEViKA's expansion logic) had the project selected those blocks for expansion.

The details on the questionnaire and selection of villages to survey are discussed at greater lengths in the following section; we pay attention to understand if the selected villages were indeed good counterfactuals on average, since the validity of the study rests on making a credible case that had JEEViKA expanded into another block, surveyed control villages had a good chance of being treated. We subsequently use the method of propensity score matching to match the treated primary sampling units (households from treated villages) to the appropriate counterparts from control areas.

3. Data & Identification Strategy

Multiple discussions with the JEEViKA team revealed that project personnel considered the Census 2001 data to identify villages with high populations of SC/ST, regarded as target population. Such villages would always get the highest priority for intervention. Grassroots personnel would then enter the village and identify the hamlets where the SC/ST populations live. The spearhead team from the project would then hold a meeting in the center of such

hamlets and inform the villagers about the project, the benefits of regular saving and arrange an exposure visit to a project village. Mobilization would start when 10-15 women from such communities commit to a weekly savings amount and federate themselves into an SHG.

The discussions with the JEEViKA team pointed out that for each block, prioritizing villages for entry was contingent on the number of total households & target (or low-caste) households in the village, as per Census 2001. Once the block-level plan had been formalized and the sequence of village entry finalized, the field team would conduct some initial scoping to look at the priority villages more closely. Specifically, they would consider the number of women in the village who are functionally literate, as JEEViKA mobilizes community members to perform as book-keepers and act as resource personnel to handhold the community institutions of SHGs and VOs. Additionally, the scoping team would also look at the number of people who are working in the village or locally; this information would be helpful when the VO becomes mature enough to conduct the interventions for various livelihood options.

In light of these discussions, the research team considered village level data from Census 2001 in 18 administrative blocks across 6 districts of Bihar, namely, Gaya, Khagaria, Madhubani, Muzaffarpur, Nalanda & Purnea. Out of these 18 blocks, 12 blocks were marked for the JEEViKA program in October 2007. Field operations in 5 of the remaining 6 blocks had started in early 2007. The remaining block, Bochaha in Muzaffarpur, was the pilot block for this program and field work had started here in late 2006.

In these 18 blocks, the research team considered 200 villages that were entered by the JEEViKA project at various points during 2008. For the purposes of this study, these villages were considered as the treatment units and all surveyed households in a treated village were considered beneficiaries of the JEEViKA program.

To look for counterfactuals, we consider villages in a separate set of 21 blocks in 5 of these 6 districts (excluding Khagaria). When the retrospective survey instrument was administered in early 2011, the JEEViKA project had just brought these blocks under its ambit; the block management offices had been set up and some initial scoping had been done to understand the logistics behind future interventions. After the retrospective survey was completed, the project scaled into 26 blocks, including *all the 21 blocks containing the control villages*.

To identify the proper counterfactuals for the 200 treatment units, we consider village level data from Census 2001. The details on the variables that were used to match villages are provided in Table 3.1.

Table 3.1: Variables used to match villages (Data Source: Census of India, 2001)

Number of Households in Village	
Total Population in Village	
SC Population in Village	Information considered to compare a non-project village to a project village came from the Census 2001 dataset for Bihar. Attention was restricted to only those non-project villages of 21 blocks in districts Gaya, Purnia, Madhubani, Muzaffarpur and Nalanda. The variables provided to the left are Census 2001 village level data that were used to construct the matched sample.
ST Population in Village	
Percent Females Literate in Village	
Percent Population Working in Village	
Percent Workers Main Workers in Village	
Percent Females Working in Village	
Percent Working Females Main Workers in Village	

The hope behind this matching was to construct a set of non-project villages from the 21 non-project blocks, which were reasonably similar to the set of project villages from the 18 project blocks. However, there is a potential problem that may invalidate this 'reasonable similarity'. Recall that JEEViKA targeted villages (in the 18 blocks) for entry based on data from Census 2001; once the village was scoped in 2008, it is possible that the field personnel found out that due to migration, the caste profile of the village had changed. This creates the possibility that the project would change the intensity of mobilizations drastically, especially given scarcity of resources at its disposal. We have the potential of a bad match if a village that is selected as a counterfactual unit, on the basis of 2001 data, does not retain the required demographics for JEEViKA to intervene in 2008.

To address such issues, the survey was administered to 10 randomly selected households from the target hamlets in all 200 project and 200 non-project villages; we can assume that had caste compositions changed significantly since 2001 in either the selected project or non-project villages, this should be reflected in the sample statistics. It is to be noted that the survey team did not have a beneficiary list for the treatment villages; thus the selection of interviewed HHs were truly random, and not a sample of beneficiary HHs only. An identical survey instrument covering several broad areas on socio-economic indicators was administered to each of the 4000

households. The instrument had two broad modules; the general module was administered to a responsible adult (preferably HH head), and the women's module was administered to an ever married adult woman. The general module collected economic information focused on asset ownership, debt portfolio, land holdings, savings habit and food security condition; social indicators attempting to capture changes in women's empowerment focused on women's mobility, decision making and networks were part of the women's module. The demographic profile of each household was captured by an appropriate household roster and caste-religion profile; in addition, a livelihood roster was also administered. Given the retrospective nature of the study, questions on certain indicators were designed to capture the levels at end 2007, along with the current level. However for other indicators, like debt portfolio, questions for end 2007 levels were not asked since the chances for incorrect responses are considerable.

The first agenda is to check for balance in treatment and comparison groups on dimensions which are invariant to interventions, but which may interact with interventions to cause impacts. To start the procedure of checking for balance in key variables, a distinction needs to be made to identify which variables are relevant for analysis at the individual level, and which are relevant for analysis at the village level.

Balance in key variables at village level enables an answer to the question: **If the project had gone to control Village B instead of Treatment Village A, could we expect to see similar impacts?** Now a similarity (difference) in impacts could be due to a combination of several characteristics in the village, and how the characteristics interact with the project, once it enters. Thus it is important to understand whether the village characteristics are similar, and whether the project interventions would have been similar in the villages. Note that the answer to this question is of paramount importance when we construct the counterfactuals; after all, if we cannot reasonably infer that Village B would have been intervened if JEEViKA went to that relevant block, then it is not very useful to consider households from village B to construct counterfactuals. We carefully examine sample characteristics at the village level to understand if the 200 non-project villages are a reasonable image for the 200 project villages.

a) Balance in indicator variables determining project expansion

We look at the determinants of project expansion first. At every level of the project, officials are given macro targets like achieving an N number of SHGs and X number of SC/ST beneficiaries. Under such targets it is optimal for the project to roll out into

a) Villages which have high levels of target population to raise chances of meeting the joint target levels, N SHGs and X SC/ST members.

b) Villages which have high proportions of target population in smaller villages to raise the chances of enrolling X SC/ST members.

c) Larger villages, but maybe smaller numbers in target population, to raise chances of forming N SHGs.

The choice is clear: Rolling out in (a) type villages is better than the other types. However the choice between (b) and (c) is fuzzy. Assume in late 2007, that instead of Phase-1 (actually entered) Block A, the project had decided to roll out in Phase-2 Block B (entered in late 2010), where both blocks are in the same district. Consider that identical targets were provided whether the block in question was A or B. **Would the project manager follow the same strategy for expansion in the control villages that he had followed for the treated villages?** With reasonable confidence, the answer is Yes, if the project manager faced similar distributions in levels of target populations and total households in both blocks. We can also consider a related question: **could a similar target be feasible in both blocks?** Once again, the answer is Yes, if the blocks in question had similar number of villages with similar distributions of target populations.

Thus the first checkpoint for balance is to identify if the control villages match up to the treatment villages in terms of the distribution of the above variables. When the project was operational in the first 18 blocks, targets and strategies were based on data from Census India 2001. The strategy for balance checks thus relies on the Census 2001 dataset; the total target population (SC+ST) is calculated in each village. The overall distribution of the Target populations in the 400 villages is considered, which provides us with mean and standard deviation of the distribution. Each Standard Deviation interval is considered as a stratum.

Villages are then grouped into strata based on their target population level. We then need to check if across each stratum, similar numbers of treatment and control villages are present & if the total and target populations are similar in each stratum across treatment and control villages.

Table 3.2: Distribution of project and non-project villages across strata of target population

	STATUS		
	Non-Project	**Project**	**Total**
Stratum			
1	122	116	238
2	57	55	112
3	13	14	27
4	7	7	14
5	1	8	9
Total	**200**	**200**	**400**

H_0: Distribution of villages is similar across status of intervention: p-value (Chi-square) = 0.225

Table 3.3: Distribution target population (low caste) and total number of HHs, by status of intervention, across strata of target population

		Distribution of target population			Distribution of total no. of HHs		
		STATUS			STATUS		
		Non-Project	**Project**	**p-value**	**Non-Project**	**Project**	**p-value**
Stratum							
1	Mean	326.6	297.3	**0.2101**	229.5	250	**0.5088**
	S.D	177.3	182.9		22.6	21.1	
2	Mean	949.7	920.8	**0.3901**	715	620.5	**0.2948**
	S.D	22.7	24.6		76.7	45.1	
3	Mean	1586.7	1619.2	**0.6788**	1455.5	1233.9	**0.5154**
	S.D	49.4	59		310.5	147.6	
4	Mean	2264.3	2345.4	**0.5511**	1713.6	1357.4	**0.1462**
	S.D	87.3	99.4		219	67.6	
5	Mean	2668	3287.1	**NA**	3279	1801	**NA**
	S.D	NA	160.5		NA	276	

Table 3.2 reveals that the number of villages by each strata of target population (apart from Strata 5) is statistically similar across project and non-project areas. Table 3.3 implies that in these villages the number of households affiliated to low castes and the total number of households was statistically similar across status of intervention, for each stratum. Together, they imply that similar targets were possible had the project rolled into the non-intervened 21 blocks, instead of the actually intervened 18 blocks. Not only that, the similarity of the numbers of target population and total households imply that block project managers would follow a similar expansion strategy in either case; distribution of villages of type (a), (b) and (c) is similar in the intervened 18 blocks vis-à-vis the non-intervened 21 blocks.

b) Balance in indicator variables for village quality

It can be argued that even with similar intensity of expansion in villages across status of intervention, village quality may have an important say in the manifestation of impacts; after all, a village with better infrastructure might be paid more attention by project staff, as mobilization in such areas makes their job easier. On the other hand, due to geographical and economic segregation, villages with better infrastructure might have little or no populations of low castes. Thus, they may not be on the radar of JEEViKA at all. Although there may be ad infinitum indicators of village quality, we consider the presence of three key public amenities at the village level to identify if treated and control villages are similar, at least in the existence of these three amenities. The three indicators considered are the presence of a school, a PDS (Ration Shop) and a Primary Health Center in each village.

Table 3.4: Distribution of percentage of villages without given amenity, across status of intervention

		Non-Project	Project	p-value
Situation of Amenity				
School Absent in village	Mean	0.07	0.085	**0.5748**
	S.D	0.018		
PDS Absent in village	Mean	0.32	0.33	**0.8309**
	S.D	0.033	0.033	
Health Center Absent in village	Mean	0.61	0.585	**0.6102**
	S.D	0.034	0.035	

Tables 3.2, 3.3 and 3.4 prove that on the basis of available data, coupled with an understanding of the expansion strategies of JEEViKA, we can claim with substantial confidence that the grassroots managers would have faced,

a) Similar targets

b) Similar distribution of target population and total population in villages

c) Similar basic quality of villages

in the 21 blocks had they been intervened in the first place, instead of the actual 18 intervened blocks. This is a key result; we can now use matching techniques to look for counterfactual households from the non-intervened villages for the beneficiary households in the project villages. Constructing a counterfactual is not a useful exercise if the average non-project village in question is radically different from the average project village, since chances are that the former village would not have been intervened by JEEViKA in any case. The above results nullify such a scenario.

We are now in a position to consider techniques for appropriate construction of comparison units; we use matching methods through propensity scores for this. As with all PSM based studies, the choice of variables that are used to generate the propensity score assume considerable importance. We now combine the thoughts from existing work in this area with knowledge of the project to identify the candidate variables that should be used to generate the propensity scores.

Let a population of N units be divided into two sets of n_1 and n_2. Let a representative unit from each set be denoted by i_1 and i_2 respectively. Let an intervention T be administered to the units in set n_1. Heckman (1997) pointed out that the relevant statistic is the ATT (Average Treatment Effect on Treated) to measure the success (or failure) of the program and is given by

$$E(\Delta Y|T) = E[Y_{i1}|T = 1] - E[Y_{i1}|T = 0]$$

The problem of the missing counterfactual is that the 2nd term is not observed. Experimental studies approximate the 2nd term by randomization; hence if the population units were assigned to sets of n_1 and n_2 randomly, the effect of treatment could be consistently estimated by

$$E(\Delta Y|T) = E[Y_{i1}|T = 1] - E[Y_{i2}|T = 0]$$

However if separation into the sets was by some rule, then the above expression is an inconsistent estimate of the ATT, since the units i_1 and i_2 are fundamentally different from each other.

Rosenbaum and Rubin (1983), Heckman and Robb (1985) and Lechner (1999) proposed a quasi-experimental approach to exploit knowledge about assignment of treatment to properly identify the control units from the set n_2 for the beneficiary units in set n_1. The essence of this approach is to note that if we can observe the levels of variables which affected the assignment of treatment, then if we can find a pair of units (one from each set) with the same levels on the same variables, either unit is the counterfactual of the other. This known as the Conditional Independence Assumption, which essentially proposes that if assignment of Treatment was a function of a vector of covariates, that is, $T = f(X)$ then

$$Y_{i1}, Y_{i2} \amalg T \mid X \quad \text{where the symbol } \amalg \text{ denotes independence}$$

In such a case, the ATT can be consistently estimated by $E(\Delta Y | T) = E[Y_{i1} | T = 1] - E[Y_{i2} | T = 0]$

Note that the vector of covariates X affects treatment, but not the other way round; for example consider a poverty reduction program which targets beneficiaries after conducting a baseline survey to identify the households below a certain poverty line. The vector of covariates would then contain the consumption levels, asset positions and other poverty indicators; however they must be measured at pre-treatment levels (for both treated and control units) to construct counterfactuals. Of course, time invariant variables (like caste) which contain information about poverty and hence influence treatment assignment should also be included in the vector X.

Constructing matched pairs for a given value of X becomes improbable when the vector has multiple dimensions, and is complicated even more by continuous elements in the vector. Rosenbaum and Rubin (1983) showed that a balancing score, b(X) which is essentially a scalar projection of the vector can be of substantial use to redress this 'curse of dimensionality'; indeed, if potential outcomes are conditionally independent of treatment assignment given the vector X, they are also independent of treatment assignment given the index b(X).

The propensity score p(**X**), which is essentially the probability of treatment as predicted by the vector of regressors **X**, is an excellent candidate for the balancing score; matching on the propensity score allows the proper construction of the counterfactual Y_{i2}, which allows us to estimate the ATT.

We now consider the broad types of information that we use to construct the propensity scores. The 1st category consists of household level variables which cannot be affected by the project, but may interact with interventions to cause differential impacts. For clarity, such variables are regarded as **time invariant variables.** For example, if education of the HH Head is systematically higher in treated areas, then one can argue that practicing financial wisdom through SHG participation would have a greater impact in treated areas. The problem is that in that case it would be tricky to ascribe what part of the impact is due to higher education, and what part is due to the intervention. Note that in various econometric settings this is still feasible, especially since the AFC data collects the information of the HH head. However we are in trouble when we consider the fact that higher education probably indicates higher motivation and abilities, which are not collected in the data (or in any data set for that matter). In such a scenario, it is impossible to ascertain what part of the impact was due to a) higher education in treated areas b) highly motivated individuals in treated areas and c) just due to the intervention itself.

The above discussion motivates why one needs to first check for balance on time invariant characteristics. This brings us to the 2nd category of household level variables on which balance checks are necessary. Consider an indicator for project impact, for example, the number of cows in a household in 2010. If treated households systematically had a higher number of cows in 2007 than control households, then comparing the 2010 levels would overestimate the effect of the project in increasing the holdings of cow. On the other hand, if control households had systematically higher holdings in 2007 than treated households, then a comparison of 2010 levels would underestimate the impact of the project. Thus, a balance check is necessary on the **pre-intervention levels of outcome variables** before one gets into discussing impacts.

Note that in case balance does not exist (for one or both categories of variables), a comparison is not impossible; attention has to be restricted to those treated and control households which have similar levels of indicators. Various matching strategies can be employed to identify units to

which attention should be restricted to; but more on that later. Of course, the village level indicator variables on amenities and target population levels are included in the balancing analysis. The detailed list is provided in Table A3.1, A3.2 and A3.3 in the appendix.

These variables are used in a probit specification, where the dummy indicating whether the observation in question is a treatment or control unit is the dependent variable. The predicted probability of participation is the propensity score, and is used in conjunction with various matching methods to generate the counterfactuals.

Some words about the specifications that are used to study the impacts are in order here; although the score generating mechanism is always a probit specification, we consider two broad cuts of the data, each of which have two specifications. The details are as follows;

Spec 1a) All households with complete information are considered in the analysis; however only economic outcomes are under study.

Spec 1b) Around 90 households did not provide information on the women's module, and 90% of such observations came from control areas. To look at all outcomes (economic + empowerment), we repeat the p-score estimation and matching algorithms to construct the ATT for all households with complete information from general and woman's module.

Spec 2a) Some of the surveyed households did not have any outstanding loans; since the most basic intervention of JEEViKA is to provide micro-credit, it would be instructive to consider the debt portfolio of the households. To do this, we consider only indebted households in this specification, rerun the complete analysis and consider only economic outcomes.

Spec 2b) In this last specification, we consider indebted households which provided information in both general and women's modules; thus, we are in a position to look at all economic and empowerment changes across indebted households in this specification.

A potential stumbling block to this study is in the retrospective nature of the instrument, which in turns raises the potential of recall error. Usually, there is no clear reason for a recall error to have a different character in general across treated and control groups. But consider an outcome which might change substantially, and change at a quicker pace, due to interventions. For example, field experience reveals that a member experiences increased freedom to move within 3-4

months of joining an SHG. Now, in January 2011, when a question was asked to a beneficiary about whether she went to a particular place at the end of 2007, there is a considerable risk that she might reply yes, although that increased mobility may have materialized 6 months down the line. Recall errors on such outcomes, which can materialize in the short run, are always going to bias the outcome upward at 2007 levels due to extrapolation by the respondent.

Indeed we can consider a question to identify if this extrapolation is actually taking place. In the mobility section, the respondent is asked whether she went to SHGs during end 2007. Around 15% of the respondents in the treatment areas said that they did; however, it is a fact that there were no SHGs (run by JEEViKA) during that time, and almost none of these respondents were part of any SHG prior to their current affiliation with JEEViKA.

What might happen if outcomes, which are subject to a systematic recall error of the above type get included in the matching process? Note that by their very nature, such outcomes are going to be higher in treatment areas at 2007 levels, which means that they will have a strong and significant contribution to the estimation of the propensity score. Now consider two potential matches, identical on all dimensions apart from the outcome on recall-error prone variable vector, say, mobility. Recall errors on that vector would then imply that the estimate for the propensity score of the treated household diverges from that of the control household; the distance in p-scores contributed by the vector may invalidate an otherwise excellent match.

Thus, among variables which have 2007 levels, we have only considered those for which impacts should materialize over a longer time horizon. In fact, the only outcomes from the women's module that has been considered for balance at pre-impact levels are whether the respondent would be able to engage in collective action when faced with some issues. The reason is that collective actions can materialize when sufficient numbers of women have joined the SHG movement in a given village, and that should take a longer time to happen than say, increased mobility to a given place.

However, this opens up the analysis to a reasonable challenge that since 2007 levels are not considered on matching, ATT estimates of 2010 levels on such variables would not account for the fact that 2007 levels were *actually* different and this difference was not due to recall errors. To address this concern, all variables (for which 2007 figures are available or can be generated)

have been considered at two different specifications while constructing the ATT. The 1st specification is the level at 2010; hence the ATT is a first difference. The other level is the *Delta-Outcome*, the difference in 2010 from 2007. Hence, for variables which were not used for balancing at 2007 levels, the ATT on the delta-outcome consistently estimates the change across the groups; a caveat being that the groups did not share divergent trends during 2007 and before.

How does recall error on a variable affect its ATT on the delta-outcome? Consider a situation where there are significant recall errors on a vector, say the mobility vector, where some respondents in the treated area systematically respond that they went to different places at end 2007, when actually they did not. If the same respondents still go to these places, the delta on these observations is essentially 0. This implies that for variables prone to recall errors, the estimated ATT on the deltas will be biased downward, the bias depending on the extent of recall error. Thus to summarize, **in case a recall error causes an upward bias in 2007 outcomes in treated areas, the ATT on the Delta-outcome will be biased downward and vice-versa.** An ATT estimate would hence provide a lower bound on the actual impact.

The delta-outcome variables play another significant role. Note that the matching technique matches on propensity score, and not exact covariate matching. Thus it is completely possible that although matches have close propensity scores, they diverge on the 2007-level of some of the balancing variables. A balance check is always performed to check for significant differences in average level across the treated and control groups; however, this does not imply that the individual matched pairs are actually similar on all dimensions of pre-outcomes. To consider a crude example, imagine that a treated and a control HH have been earmarked as a match for each other, but had dissimilar holdings of, say, cows in 2007. If the 2010 level is comparable, the contribution towards the ATT would be negligible. However, the delta for the HH which increased its holdings would contribute much more towards the ATT on the delta for the overall sample. Thus, considering the delta-outcomes, along with the first difference increases the confidence in changes, as the delta controls for level differences at 2007 and just considers the net change in 3 years.

Hence, the delta-outcomes play a dual role: they mimic the advantages of a Difference-in-Difference estimation, but are able to allow information in time invariant characteristics to construct the counterfactual, when such variables are used to estimate the propensity score. Do

note that the assumption of similar trends apply to either process of estimation for consistent results.

If the 2007 level is balanced across T-C on average, then a significant ATT on the first difference will imply a significant ATT on the delta. In fact it would be a very odd result, if for outcome X, 2007 levels are balanced, 2010 levels are significantly different but the delta is statistically similar across groups.

However, if the 2007 level is not balanced across T-C on average, we may have a significant ATT on the first difference, and an insignificant ATT on the delta, which implies that the groups are moving similarly. In fact, if the ATT on the delta is positive, it can probably be said that the gap is closing.

A significant delta will not imply a significant ATT on the first difference, due to inexact covariate matching at 2007 levels. In this case a significant delta contributes towards the confidence in impacts.

To summarize the discussion on recall errors:

1) A systematic component of the recall error may bias the 2007 level of some outcomes upward in the treatment areas. Using such variables in matching would raise chances of inexact matches. Thus such variables are not used for matching. However the deltas are used, along with first differences, to address the issue that had the 2007 levels been used, ATT estimates on the first difference might be very different; the key point is that the estimated ATT on the delta, if recall error of the above kind has taken place, will be a lower bound on the actual ATT.

2) Since exact matching on all covariates at 2007 levels is impossible, the estimate on the ATT of the Delta-outcomes raises confidence in the presence or absence of impacts, as the delta removes the concern of mismatch at 2007 levels.

Hence, the broad types of variables considered:

Type A: 2007 level is available or computed. 2007 level is used for matching and balance. ATT on 2010 level and ATT on Delta are computed.

Type B: 2007 level is available or can be computed. However, 2007 level is not used for matching and balance. ATT on 2010 level and ATT on Delta are computed.

Type C: 2007 level is not available. Hence only ATT of current responses are computed. The implicit assumption is that Type C variables are highly correlated with both Type A and B variables.

Before we move on to the algorithms for matching, we briefly digress to discuss systematic recall errors that may be introduced on the account of any retrospective values. Given the previous discussion, it is clear that if beneficiaries ascribe changes in outcomes at the retrospective level, the ATT would underestimate the true effect. It might be argued that beneficiaries may underestimate pre-treatment outcomes and paint a 'worse' picture than it actually was, before the program came in. This might be due to a psychological effect of imagining a worse situation than it actually was; it may also be due to a strategic ploy on part of beneficiaries to paint a better picture about the program. This would be a sensible ploy only if the beneficiaries know that the program is being evaluated and they have found the program actually beneficial. A counter-argument may be that under such a scenario, beneficiaries may under-report current outcomes, if they assume that reduction in poverty may remove them from the program's ambit.

In any case, if a systematic recall error causes beneficiaries to underreport retrospective levels, the difference in outcomes at current periods would overestimate the actual effect. If under this situation, beneficiaries underreport current levels, then there is a downward bias. In any case, the absence of a true baseline complicates our understanding about the direction of bias if systematic recall errors exist. Indeed, the data points out clearly that on some dimensions, beneficiaries are ascribing program outcomes to retrospective scenarios; for example, claiming that they did go to SHGs when it is a fact that SHGs did not exist. We know that under this scenario, ATTs on the current outcomes are a lower bound on the actual effect. However, a-priori we do not know which outcomes are subject to systematic recall errors, and in what direction. For this reason, we re-run Specifications 1b and 2b without any outcome variables measured at retrospective levels. The results on balance and subsequent matching from these re-runs are presented in the appendix, as an additional robustness check on the main specifications, which still include the retrospective levels of outcomes.

We are now at a stage to discuss the various matching protocols that are used in the current study; 5 matching methods have been used to construct the counterfactuals. The 1st two methods are NN (with replacement) matching and kernel matching, where the bandwidth is given by the auto-generated rule of thumb optimum. The 3rd method is also a kernel algorithm; it uses a bandwidth which comes out of minimizing the root mean square error (RMSE) by using a process of leave one out cross validation (LOOCV). The Leave-One-Out-Cross-Validation (LOOCV) process uses a minimization criterion of the RMSE to identify a reasonable bandwidth. The last 2 methods considered are a caliper and radius specification with the same tolerance level. We recall that the choice of this tolerance level is important for caliper/radius specifications; hence, we spend some time to discuss the reason behind choosing the tolerance level, which in the present study is given by:

Tolerance Level=   (SE of Average Treatment Probability of Treated Observations) –

(SE of Average Treatment Probability of Control Observations)

We start by looking at the estimation of the propensity scores and their distribution among the treatment and control units; these are distributions are from the unmatched sample.

Figure 3.1: Distribution of Propensity Scores, by Intervention Status, across Specifications

(Distribution of propensity scores of Treatment & Control units in green and red respectively)
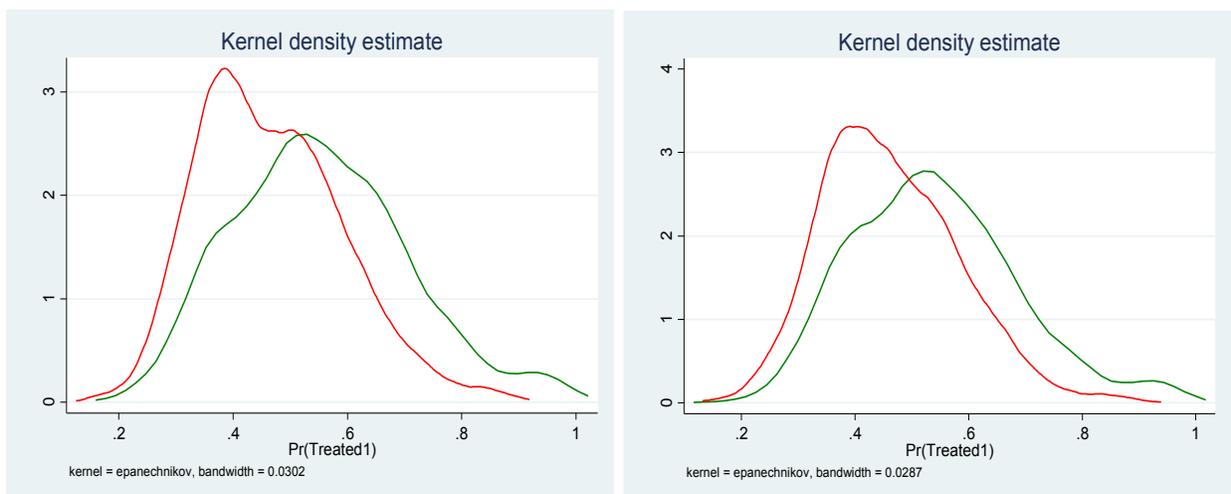
Spec 1a: All HHs, only variables from gen module used   Spec 1b: All HHs, variables from both modules used
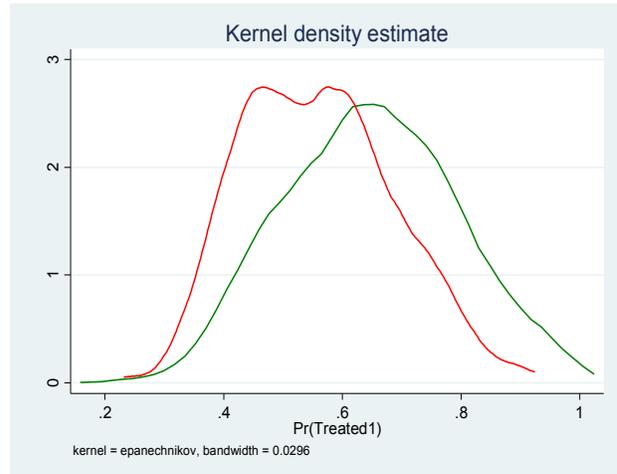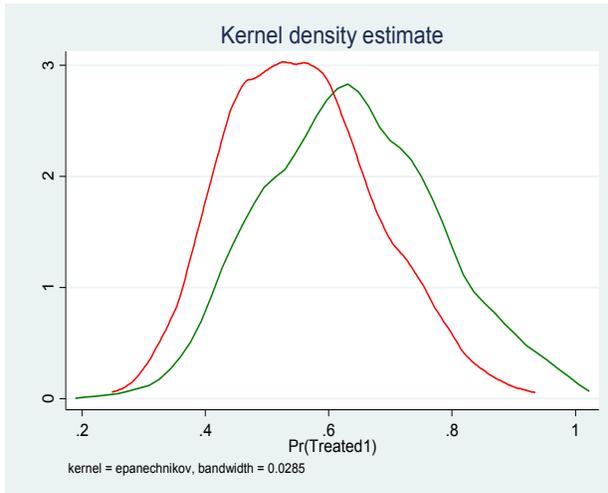
Spec 2a: Indebted HHs, only variables from gen module used   Spec 2b: Indebted HHs, variables from both modules used

The distributional graphs contain a major implication; a substantial number of observations from either treatment or control sets are in the common support region. Below, we provide the graphs of distribution of matching and the statistics on post-match balance for Spec 1a to understand the intuition.

Figure 3.2: Distribution of Matched Units, across Match Algorithms
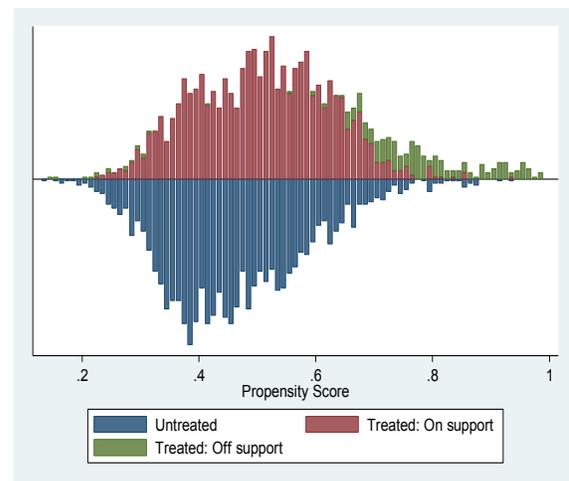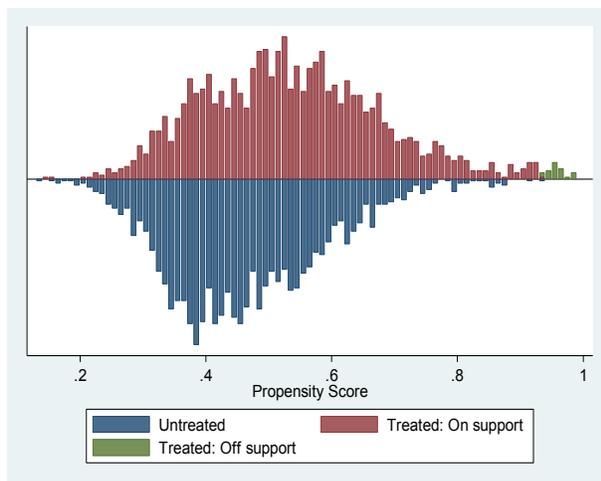
Nearest Neighbor Post Match Graph   Radius/Caliper Post Match Graph
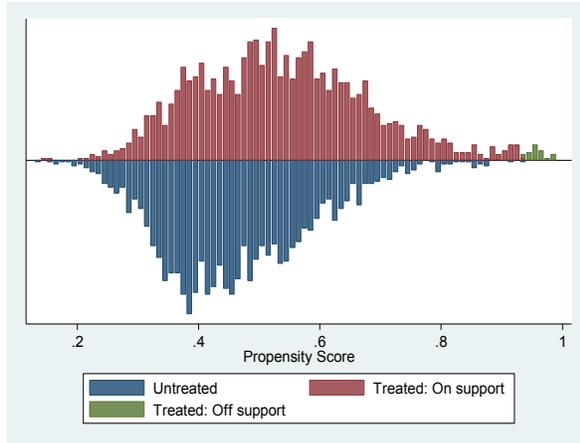
Red: Matched Treated Units

Green: Unmatched Treated Units

Blue: Control Units

Kernel Post Match Graph

In the balancing exercise, common support had been imposed; this essentially means that the treated units with a propensity score higher than the propensity score of the control unit, with the maximum propensity score, are not considered for matching. For nearest neighbor and kernel algorithms, this is the implication of common support. Note that in nearest neighbor and kernel, all treatment units are matched; additionally, in kernel matching, all control units are used to construct the match. In radius/caliper algorithms, the imposition of a tolerance bound, say ε, implies that all treated units which do not have a control unit within a distance of $|\varepsilon|$ in propensity scores are left unmatched. Thus under radius/caliper algorithms, the quality of matching (in terms of proximity of propensity scores) is decreasing in ε. In table A3.4 we look at the balance statistics on the pre-treatment levels of the outcome variables for Spec 1a.

We are now in a position to interpret the results. Due to the number of specifications, algorithms and probable outcomes, we have a large set of ATTs to consider. In the following discussion we focus on the results that are generally robust, especially when we consider specifications 1b and 2b. The detailed results across specifications and matching modules are provided in the appendix.

4. Results

We first look at outcomes on livelihoods, keeping in mind that the survey instrument was not geared towards pinpointing changes in incomes from various sources due to the retrospective nature. Instead, we try to understand if such changes happened by considering a variety of proxy

indicators, such as number of income earners, substitution among livelihood activities, land holding and leasing patterns and finally, buildup of assets.

4.1 Impact on Livelihood Options

JEEViKA was unable to change the **number of income earners in beneficiary households**, irrespective of the income being seasonal or year round. Now, this may not signify absence of impacts once we recognize that JEEViKA does not provide employment opportunities, but attempts to expand livelihood options (an avenue of generating income). Thus, income earners in the beneficiary household may either allocate more time to their existing livelihood(s) or substitute towards a better livelihood option generating higher net income. Due to the retrospective nature of the instrument and the difficulty in collecting accurate income figures from rural India, we do not ask for income earned for each past and present livelihood. Rather, we look at the livelihood activities (by season) and attempt to infer something from that. The results on the shifts away or toward a particular livelihood option are generally non-robust, and small. However, there is a robust result for animal husbandry as an option; **0.5 % treated households are shifting towards animal husbandry as a primary livelihood option**, across most algorithms and all specifications.

4.2 Impact on Assets

Current ownership of mobiles in 2010 is significantly higher among treatment households (56%) as opposed to control households (51%), under almost all algorithms and specifications. Although there are some positive results on change in holdings of other assets like watches, they are small and non-robust. No effects could be found on land ownership or leasing in. However, there are a couple of results, which are interesting in the light of a previous result on the livelihoods options.

When we consider indebted households only, the **holding of buffaloes in control areas (6.2%) have accelerated faster than that in the treatment areas** (3.7%) over the past 3 years. However, when we consider all households**, treatment areas (4.6%) increasing their holdings of cows over the last 3 years as against the control areas (2.9%).** We recall a result from the discussion on the livelihoods dimension; a small, but sure shift towards animal husbandry. The confusing part is that ownership levels over the last 3 years are moving in opposite directions for

26

cows and buffaloes, which are essentially substitutes. The buffalo is monetarily dearer than the cow, and provides better milk, but the cow brings an immense value of prestige and sentiment with it. The instrument does not collect any details on the leasing in of animals, which is a substantial activity under animal husbandry. Future work may provide a better understanding of this result.

4.3 Quality of Housing and Food Security

Across all specifications, but usually under radius algorithm, there is evidence that the **percentage of households with flooring made of permanent** or 'pucca' (cement, concrete, etc.) **materials in the house has increased at a faster pace in control areas(1%) than in treatment areas (0.5%).**

A small promising effect materializes across the board when we consider the defecation practices. **In the past 3 years a significantly higher proportion of treated HHs (3%) has stopped using open fields for defecation, as opposed to control households (1.5%).** Use of closed public or private toilets has increased in project areas. Indeed, **around 1.6% households from treated areas have started using private toilet facilities over the last 3 years**, **compared to 0.8% control households.**

However, we need to put this change in perspective; a high percentage of the population (around 86-90%) still use open fields for defecation in the present day, both from project and non-project areas. A lot of work remains to be done in this area, given the fact that defecation in the open leads to a plethora of health problems.

Across specifications and under kernel algorithms, there is evidence that the duration of acute **food shortage has reduced over the last 3 years in treated units**. **However*, this reduction is extremely small, although significant and robust at about .09 months.* Once again, the absolute number of months of acute shortage is very high (around 1 month for the control areas, 27 days for the treated areas) which is why the difference works out to around 3 days.

Across specifications, especially for kernel and radius algorithms, **there is evidence that the percentage of insecure HHs has reduced faster over the last 3 years in project areas**. This effect is to the tune of 2.1-2.9 more households from treated areas, per 100 HHs from either area.

4.4 Children and Woman's Profile

The **enrollment figures for the girl child are significantly higher in treated areas, under Spec 1a and the NN, caliper and radius algorithms.** Around 8%-10% more girls attend schools in treatment areas. However, these results are not repeated under other specifications. The **enrollment figures for the boy child are more significant, for the indebted households, apart from the caliper algorithm.** Around 8%-13% more boys are currently enrolled in schools from the treated units.

Respondents from both treated and control areas wanted to marry off their daughter when she is 16 years old; and due to this, there were no differences along this outcome across specifications or algorithms. The NN and caliper algorithms under Spec 1b imply that **women from treated households want to educate their daughter for 0.4 years more on average**. However, the significance is lost for the other specification as well as the other algorithms.

Women from treated areas seem to be much more interested in their son's education; **women from beneficiary HHs want their boys to be educated for 0.47-0.54 extra years**; when we focus on indebted HHs, women want to educate their sons for 0.42-0.54 more years.

33%-34% more women are signature literate from treated areas. Now this is, to a large extent, a trivial impact. Women are encouraged to sign their names in JEEViKA SHGs. We can consider the ATTs on sign literacy to understand if women are getting keener in recognizing numbers or letters. **Around 3.3-4.4% more women are sign literate from treated areas** under Spec 1b. Once again, scope exists in this area as percentages of sign literacy are in the range of 16-20% in the entire sample. Lastly, we consider the percentage of women who mentioned their husband's name by themselves. Although this is no direct indicator of empowerment or well-being, orthodox societies consider this as taboo. It is interesting to see that 15-17% more women from project areas do not view it as such.

Up to this point, we considered results on assets, livelihood options, house quality and food security. It is worth noting that apart from the last dimension, positive results on the other dimensions would probably indicate that the household has come out of poverty. Clearly, we do not have extensive results on these dimensions; however, the small sporadic effects indicate that

the direction of change is optimistic. We now consider 2 of the 3 main thrust areas of JEEViKA, that of micro-finance and women's empowerment.

4.5 Savings Habits and Debt Portfolio

We note that JEEViKA members are highly encouraged (in fact, required) to deposit a weekly saving in their Self Help Group. Thus an impact on savings is expected. We consider the regularity of savings at current levels and changes in such behavior over the last 3 years; additionally, we consider where these savings are usually parked. **95% households from treatment areas practice regular savings currently, as opposed to 24% households from control areas. Around 70% treatment HHs started regular saving over the last 3 years, compared to 12% control areas.** Quite obviously, SHGs have become the dominant place to park these savings, at the cost of non-formal and other formal mechanisms.

Although these impacts are structural (robust across specifications and match modules) and simply massive, we should note that this is somewhat trivial. A more fundamental change would have been had beneficiary households saved larger and larger amounts voluntarily. Unfortunately, the instrument did not probe for voluntary saving amounts (rather, any savings amounts) due to the retrospective nature and the fact that a concept of voluntary savings is confusing in non-SHG areas. However, we need to take cognizance of the fact that even a token savings practice is absent in impoverished households of rural Bihar. At the end of the day, weekly savings to the tune of 5-10 Rupees is still an achievement, given the resource constraints on JEEViKA's target population.

We now consider the (and perhaps the most important) dimension of debt portfolio. The pernicious poverty levels in rural Bihar are engendered to a large extent by high cost informal credit markets, and complete unavailability of formal credit. Emergency situations make expensive credit unavoidable, leaving fewer resources to take credit for productive purposes. Assets get mortgaged, leading to the debt trap; the extent of the debt trap leads to occurrences of bonded labor in some areas. We take a careful look at the debt portfolios to understand whether JEEViKA has been able to crack this problem at all.

The retrospective nature of the instrument meant that we could not look at the initial credit position of any household. Indeed, the best indicator for historical indebtedness is the year of borrowing; one could look at the amounts and purposes of old loans (that are still outstanding) and make some inferences. This is exactly what we exploit, by looking at loans taken on or before 2007 and since 2008. Note that loans taken on or before 2007 are not variables that we should balance on; if the intervention takes root, old loans should get retired much faster in treated areas. Hence, we cannot balance on any debt related variables.

We take a close look at the distribution of high cost (monthly interest rate greater than 2%) loans, separated by the year of 2008, across treated and control areas. We then look at loan uptake by purpose. Immediately we run into a problem; interpretation of amounts borrowed by purpose doesn't make a lot of sense if we cannot control for the entire portfolio. For example, cheap credit may encourage loans for consumption and/or productive purposes in treatment areas. However, if beneficiary households keep using credit for consumption purposes, then the beneficiary households may be getting to higher credit equilibrium for the time being, but that's about it; such practices would not translate into higher incomes. Just a casual comparison of the absolute borrowing by purpose might be very misleading, as the total portfolio (and the part allotted to consumption) may be higher in project areas just due to easier and cheaper credit. Thus we need to consider percentages. This means that we necessarily consider only the currently indebted households; and this is the main motivation for Spec 2a and 2b; before we consider the structure of the debt portfolio from indebted households we consider the direction and order of the size of the portfolio from all households.

1.5-2% less households from project areas has high cost loans which were taken before 2008. Note that in any case, **5% households from control areas still have outstanding amounts on such loans**. The amounts borrowed on such loans are similar across treated and control units. Strong results show up when we consider high cost loans taken on or after 2008; program areas show a clear substitution away from such loans. **44% households from control areas have outstanding amounts on high cost loans taken after 2007, compared to 24% treatment households which are still under such high cost debts.** The amounts borrowed on such loans are **Rs 3500-4100 less in project areas; the average control HH borrowed Rs**

7750-8300 on high cost loans. Additionally, the average number of loans (any loan) is 0.5 units higher in treated areas.

The results are expected and encouraging; JEEViKA beneficiaries have retired old loans at a faster pace; significantly lower numbers of project households have taken high cost loans after the project expanded into the blocks. Additionally, the amounts taken out on such loans are significantly lower in project areas. Due to the lower cost of the loans, beneficiaries are taking more loans on average; however, the total amounts borrowed are not different.

We glance quickly at the borrowings by purpose; project beneficiaries are taking loans more frequently for a variety of purposes, including repair of house, purchase of food, marriage expenses, durables purchase, debt reduction, livestock purchase and petty business. However, the differences in amounts borrowed by purpose are not significant across the board. This implies that beneficiaries are taking out loans more frequently; however, this does not translate into higher total borrowing, implying that smaller amounts are borrowed more frequently by beneficiaries.

We now turn towards the indebted households; results on the debt portfolio become more pronounced and more clarified now, as we have the luxury of considering percentages.

Among currently indebted households, **4.9-6% less households from project areas still have positive outstanding amounts on old high cost loans. About 10% households in the control areas still retain such debts.** Once again, the amounts borrowed on such loans are still statistically similar. **47-50% less households from project areas have taken high cost loans after 2007; indeed, the percentages of control HHs with 'new' high cost debt burden is 77-80%.** The results on amounts borrowed are even starker. **Program HHs have taken around 9300-10000 Rs less on high cost loans after 2007; the control HHs borrowed around 14200-15000 Rs on high cost loans after 2007.**

Two results follow immediately, which are extremely encouraging when taken together; **indebted HHs in project areas has taken 0.18 more loans (any loans) than indebted control HHs**. However, although the number of loans is thus higher in program areas, **control units have a higher total borrowing to the tune of 5400-6500 Rs. Indeed, their total borrowing is around Rs 19500-20500.**

These two results imply that project HHs take more frequent loans, but borrow smaller amounts on each loan. This may lead to a potentially healthy practice of repeat doses of credit, if it's done for income generation activities. We also note that due to the practice of mortgaging assets while accessing loans from informal spheres, rural families usually borrow a high amount of money for multiple purposes, in lieu of mortgaging a single item. Obviously, this is a prime recipe for debt trap. Cheap credit with no requirement of mortgages has been able to crack this conundrum, and thus program families can now go for repeat doses of smaller credit. We now look at the purposes of borrowing to understand if these repeat doses are being used for short run benefits like consumption purposes.

The radius algorithms point out that there is a reduction in the number and amount of loans taken out for health purposes. **For every 100 Rs borrowed, program households take 4.4-5 Rs less for health reasons. This is the by far the most important purpose of credit in rural Bihar; out of 100 Rs borrowed by the control unit, almost 41 Rs is for a health reason.** There is a strong result when we consider loans taken for **marital expenses; program HHs have taken out 1700-2900 Rs less than control HHs** for this reason. This translates into 10-13 Rs difference, when we consider a project and control HH with a total debt of 100 Rs. **The average control HH borrows 23-24 Rs for marriage expenses, out of every 100 Rs it borrows**. Distribution of borrowing patterns is very similar when it comes to the purposes of food requirement and schooling across treatment and control areas across all algorithms. The average **indebted control HH borrows around 1900-2500 Rs for house repairs; the treated HH borrows around 700-1100 Rs less for this reason.** However, there are no significant differences in the percentage borrowed for house repair. There is some sporadic evidence of program HHs borrowing lower amounts for purchase of durables, under the NN, caliper and radius algorithms to the tune of 470-550 Rs; once again, there is no significant difference when we consider the percentage of total money borrowed for purchase of durables across treatment and control.

An extremely strong result shows up when we consider the purpose of debt reduction; **indebted program HHs borrow, on average, 700-800 Rs more than the indebted control HHs to reduce other debts.** If we **consider a program and control HH with a total debt of 100 Rs, the average control HH borrowed Rs 27-70 for debt reduction; the program HH allocates**

Rs 7 more to this purpose; in percentage terms, this is 0.48-0.80% of the total amount borrowed. Simply put, cheaper loans are being used to retire other loans by beneficiary households; it is a significant step towards coming out of the debt trap. Now, we recognize that this would be a sustainable strategy if program HHs move towards credit for income generation.

Program HHs borrow around 250-380 Rs more to purchase livestock; this translates to 4.3-4.9 Rs more being borrowed, by the beneficiary, for every 100 Rs borrowed by either beneficiary or non-beneficiary. Indeed, the average control HH allocates 1.7-2.5% of her entire portfolio to this purpose. **2.2-3 Rs more are allocated by beneficiaries towards setting up a shop, when either treated or control borrow 100 Rs total;** however, we do not see a significant difference in the amounts borrowed. Unfortunately, due to oversight, the author did not create a separate category for agriculture purpose; such information was lost as it got clubbed under others.

However, the debt portfolio sends out some very clear results; a structural change has taken place in the debt portfolio of program households. They take more frequent loans but they borrow smaller amounts on each tranche. The loan burden under high cost debt is lower by a large margin; additionally, a much lower percentage of project HHs have such debt. Old high cost loans have been retired at a faster pace by project HHs.

The results on the purposes of debt reduction, consumption and income generation point out that the borrowing pattern implies financial wisdom on the part of beneficiaries; they borrow more to reduce other debt, they borrow less to arrange for marriages, house repairs and purchase of durables. Although there is some evidence that percentages of amount borrowed are lower for health purposes, the evidence is not across the board. However, the program households have definitely moved towards credit for productive purposes, evidenced by higher uptake of loans for livestock rearing and setting up small shops.

4.6 Dimensions of Empowerment

Stable results on mobility materialize across algorithms along the entire dimension of mobility, especially when we consider all households. **We start with the results from indebted households, where the ATTs are a bit dampened**. Presently, **around 3.9-4.5% more women from project areas can go to health centers for concerns regarding themselves and their**

children. Note that the ATT on the change in mobility to health center question was not significant; there is no significant difference among treated and control areas, in the percentage of women that have started going to health centers over the last 3 years.

Before moving on, we recall the earlier discussion on recall errors; note that we had said that the Delta-outcome is of much use when exact covariate matching is not used. However, when we do not balance on 2007 levels due to the chances of recall error, the ATT on the current level may well be significant due to not balancing. However, the ATT on the Delta-outcome always provides a lower bound, and hence provides additional information towards validating impacts. An example follows.

We look at another across algorithm result from indebted households; attending panchayats meetings is really not usual among women from impoverished, low caste households. Reservation for women has not been of much use either, because the elected representative is usually remote controlled by her husband anyways. Women, especially from the program's focus households, have enjoyed abysmally low levels of voice, participation and representation in the political process at the grassroots.

However, **5.8-6.5% more women from program areas attend such meetings currently. To understand the importance of this impact, note that only 3-3.5% women from control areas go to panchayat meetings presently.** Due to the possibility of recall errors, balancing was not done on pre-impact levels of mobility. Thus we consider the Delta-outcome, the number of women who have started going to panchayat meetings over the last 3 years. The ATTs are significant (although small) on this outcome; around **1.4-1.7% more respondents from project areas have started this practice over the last 3 years**. This is where the Delta is of additional importance; clearly, it tells us that the program has made women more participatory in the political process, and that the estimated effect is still a lower bound.

We now consider the results from all households, keeping in mind the importance of the Delta-outcomes.

2.7-3.5% more women from the project areas go to neighborhood grocery stores in 2010; however the ATT is not significant when we consider the change over the last 3 years.

4.5-6.3% more women from project areas go to health centers in 2010; over the last 3 years, 2-2.6% more women have been going to this facility from the project areas. The significance of the Delta tells us that the program has been able to change mobility patterns when it concerns the health of the woman or her child.

When it comes to **visiting a friend or neighbor** in the village presently, **around 2% more women from treated areas have been able to do so; note that the percentage from control areas at present dates are at 95%.** Over the last 3 years, similar percentages of women from the program villages are visiting their neighbors. Thus the program has been able to relax restrictions on mobility, even to the woman's immediate neighborhood for the most 'strict' households.

Around **3.3-5% more women from treatment areas are able to go outside their village to visit a relative presently. Around 2-3% more women have done that over the last 3 years, from program villages.** Thus, the program has been able to engender better contacts between a woman and her networks, whether such contacts are within or outside the woman's domicile village.

The results on attending panchayat meetings presently and attending such meetings over the last 3 years are consistent and comparable with the results from the indebted households. Once again, we recognize that JEEViKA has been able to crack a very low level equilibrium and encouraged participants in the program to participate in the political and deliberative process of their community.

No significant differences can be seen on the decision making patterns with respect to daily cooking and purchase of personal items, where percentages of women who participate in such decisions are high across areas.

However, around **8-10% more women from JEEViKA villages provide an opinion in the purchase of a durable item in 2010**; around **2% more women have started doing that over the last 3 years**. To put this in perspective, 41-43% women from control areas provide their opinion on this aspect currently.

We should expect that due to the close correspondence between the outcomes "Participating in decision related to health of self or child" and "Mobility to health center for health purpose", the results should follow similar lines. This is confirmed, as the ATTs are dampened for indebted households but get pronounced when we look at all households. **2.8-4.4% more women participate in such decisions currently, while 2-2.6% more women have started participating in the last 3 years.**

Among indebted households, **10-13% more women from treated villages have an opinion when it comes to their children's education in 2010**; however, the Delta is not significant'. When we look at all households, 8.5-9.5% more households exhibit such opinions; additionally the ATT on the "Del" is significant and implies that 1.5-2% more women started providing such opinions over the last 3 years. Putting this in perspective, **68-70% women from control areas have any say in their children's education currently**.

Among indebted households, around 5% more women have a say about what and where the primary livelihood should be while 2% more women have started providing this opinion over the last 3 years. However, this result is significant only for the radius algorithm for indebted households. When it comes to her employment, 5.4-7% more women have a say presently; around 1.7-2.5% more women have participated in this decision over the last 3 years.

These results are more pronounced and stable when we look at all households together. **5-6.3% more women provide a decision about the primary livelihood activity currently;** in 2010, **6.4-7.5% more women participate in decisions regarding their own work**. **Over the last 3 years, 1.9-2% and 1.5-2% more women have started to provide an opinion about primary livelihoods and self- employment respectively**. Once again, we note that around **50% of women from control areas participate in either aspect of decision making currently.**

We'd expect that since cheap credit is coming from SHGs via the female member, the woman should have a higher say in **borrowing decisions. Among indebted households, 18-20% more women provide an opinion currently, while 6-7% have started providing an opinion over the last 3 years.** Results are stable (and more pronounced) when we look at all households. We note that currently, **only 58% women from control areas provide an opinion in the credit needs of the household.**

Lastly, we look at the patterns of decision making when it comes to politics. We recognize that having a separate political identity in such families is extremely unusual for women; a cursorily look at the profiles of some of the elected candidates from the 'Hindi Heartland' is proof enough, where irrespective of performance, caste alignments dictate political allegiance. Indeed, **from control areas, around 19-20% women have any say in the political preference of the household**. However, in 2010, **8.7-10% more women from JEEViKA areas provide an opinion in this aspect of household decision making. Around 2.6-4% more respondents have started to participate in this decision over the last 3 years.**

Once again, we see that the program has encouraged the participants to engage with her family on increased issues at increased frequencies regarding the dimension of household decision making. Such changes have materialized at either high (for example, education/health) or low (for example, politics) level equilibriums.

We now look at the propensity of program participants to engage in problem solving at the community level. We look at 4 issues, which relate to some woman being beaten up, some problem with the PDS, some problem with the school and some problem with the mukhiya (the elected panchayat chief). We ask if the woman would act if she faces such a problem, and if she does, who would she approach to take a suitable action. We focus on the percentage that would act by themselves and/or act with other women. Note that we assume that if she does not act, then she does not take anybody's help either. Thus if the woman responds that when confronted by given problem, she is unable to act, then the response for who she acts with are coded to 0. Essentially, the percentages who act by themselves or with other women are defined over the entire sample and not for the subset which says that they are capable of doing something. We should note here that 0s on 'who you would approach/act with' includes responses like 'ask my husband', 'approach mukhiya' and the responses 'can't do anything'.

Note that the percentages of women who would take some action when faced with such issues at end 2007 are used as balancing variables. Thus, although we look at the Delta-outcomes, the use of them as a triangulation mechanism for the corresponding variables at current levels is less now.

Among indebted HHs, **6.4-7.8% more women from project villages claimed that they would act if a woman is beaten up in her village in 2010; over the last 3 years, 4.2-6.8% more respondents have become capable of acting in the last 3 years. Around 5% would act by themselves and 13% (which may include some or the entire previous figure) said that they would act in sync with other women. Around 73% women from control areas would act when some other woman is beaten up in her community; however, only 36-40% of the (entire) respondents would act by themselves or in sync with other women**. The impacts are more pronounced when we consider all households, irrespective of indebtedness status.

Significant results materialize for indebted HHs only under the kernel algorithms, when we consider issues with PDS. Around **5% more women would act presently (and 5% have become capable of acting over the last 3 years) when there are issues with the PDS. 6-7% more women would work with other women in harnessing this issue.** Around **38-40% of respondents from control areas would act on this issue, and around 15% would act with other women.** When we consider all households we get comparable results and statistics on all of the above figures. There is no difference in propensities of acting by self in case of PDS related issues between treated and control areas, across specification.

Similar statistics hold true for indebted households when we consider issues with the school in the community. **5% more women would act presently and similar percentages have become capable over the last 3 years. Additionally, 5-6% more women from project areas would act with other women to resolve such issues.** The results are comparable (somewhat more pronounced) when we look at all households. **Around 36-40% of women from control areas would be capable of acting to resolve this issue.**

Women from indebted treated households are not more likely to act than their counterparts in control areas when they have some complaints against the mukhiya; however, the pattern of action is different. Around **3.4-4.7% more respondents from indebted households in program areas would work with other women to act on mukhiya related issues.**

The result becomes more interesting when we consider all households. Currently, **there is no significant difference between the percentages of women (between project and non-project areas) who would act when some issue with the mukhiya crops up.** However, **2.6-3.6% more**

women, under JEEViKA's ambit, have become capable of dealing with mukhiya related issues over the last 3 years. 4.4-4.7% more women from program areas would work with other women to resolve mukhiya related issues.

Once again, we note that this is an impact at a **low-level equilibrium; about 26-28% women from control areas said that they would be able to do anything when faced with mukhiya related issues; only 11% would be able to work with other women to try and address those issues.**

To quickly summarize the findings from the previous 3 dimensions; clearly, JEEViKA has unambiguously affected empowerment levels of women. We see clear evidence in the mobility of the average beneficiary woman, her participation in household decision making and her confidence and propensity to engage in Collective action when faced with issues related to the community. One avenue of these changes is clearly strengths in numbers; once the SHG movement spreads in the village, the woman's network inside her village keeps expanding exponentially. But these numbers progressively relax the social norms of 'right and wrong'. The detailed results on ATTs, for all households reporting information in women's modules across the key matching modules are presented in table A4 in the appendix

5. Summary

The JEEViKA program has brought about some definite changes on a variety of dimensions in the lives of the socially and economically marginalized castes in Bihar. It has freed up most of the households from high cost debt; beneficiaries have started to take steps towards using credit for productive purposes, after retiring expensive loans. The results from different matching algorithms and across specifications imply that these trends are robust and stable, as they are manifested in different cuts of the data. Additionally, beneficiaries are now practicing regular savings.

However, these trends have perhaps not translated into higher level outcomes to the extent that the project may have envisaged. For example, there is strong evidence for higher ownership of mobile phones and watches; however, there is mixed and weaker evidence when we consider clearly productive assets like cows and buffaloes. Beneficiary households seem to be increasing

the holding of cows, but reducing the holding of buffaloes. The survey instrument does not collect information on leased animals, a quite popular practice in rural Bihar.

The project seems to have reduced incidences of acute food shortage. Similarly, there are clear trends in decreased defecation in open fields. In both cases, the impacts are small in scale, and there exists scope of further reduction in food insecurity and unsanitary practices.

If we look at indicators of social achievements, JEEViKA's impacts are substantially deeper. Women aspire to educate their children more, especially the male child; we can see some evidence of those aspirations taking root, as evidenced by higher enrollments of the boy child in project areas. The ability to sign one's own name is definitely higher among beneficiaries; but if the project assumed that this would trigger an automatic interest into higher orders of literacy, then the trends show that this is not the case. There is evidence that a higher percentage of women from project areas can read numbers, letters and signs, but once again, there exists considerable scope in improving these numbers.

Mobility, decision making and propensities towards collective action can be assumed to be definite indicators of women's empowerment, especially among low caste households of rural Bihar. The program has made significant strides towards empowering women along a variety of dimensions; the results on increased political participation, an erstwhile no-go area for women, are highly encouraging.

However, it is worthwhile to note the following point; relative to counterfactuals, women from JEEViKA SHGs demonstrate significantly higher empowerment, as evidenced by a variety of indices. In absolute terms though, there is scope for higher achievements. Outcomes related to political awareness or participation are a clear example of this phenomenon; although women from project areas display considerably higher engagement in political decision-making, the scope for further changes are immense. On the dimension of participation in decision making, we find such phenomenon for every decision which relate to outcomes that are more public in nature. Thus, JEEViKA women have significantly higher say in self-employment, primary livelihoods, purchase of durables, etc., than control HHs; however, in absolute terms, the percentage of women that do participate in such decision making can still improve by a large margin.

However, we note that the retrospective nature of this survey left many questions unanswered. Although we have tried to address the problems with recall errors, we cannot be certain about the effect that they may still be playing. Finally, we recognize that such a complex, iterative package of interventions could be better measured by a package of evaluation designs, and not just a rigorously designed quantitative study. It is worth mentioning that a multi-disciplinary study is no underway at JEEViKA, which tries to understand the process, quality and quantity of socio-economic and behavioral changes among JEEViKA beneficiaries. Additionally, there are several allied randomized control trials, which measure the impacts of some standalone interventions of JEEViKA. It is hoped that the results from these allied research agendas would jointly contribute to the understanding the quality of service delivery by community driven development projects that work via institutional platforms and multi-pronged interventions to address poverty.

Appendix

Table A3.1: TIME CONSTANT VARIABLES measured at VILLAGE LEVEL

HC_Absent	Percentage of HHs in village without a HEALTH CENTER
PDS_Absent	Percentage of HHs in village without a PDS
School_Absent	Percentage of HHs in village without a SCHOOL
TargetPopLevel1	Percentage of HHs in village in Stratum 1 of Target Population [0 < (SC+ST) < 660]
TargetPopLevel2	Percentage of HHs in village in Stratum 2 of Target Population [661 < (SC+ST) < 1320]
TargetPopLevel3	Percentage of HHs in village in Stratum 3 of Target Population [1321 < (SC+ST) < 1980]
TargetPopLevel4	Percentage of HHs in village in Stratum 4 of Target Population [1981 < (SC+ST) < 2640]
TargetPopLevel5	Percentage of HHs in village in Stratum 5 of Target Population [2641 < (SC+ST)]

Table A3.2: TIME CONSTANT VARIABLES measured at HH LEVEL

Caste_SC	Percentage of SC Households
Caste_ST	Percentage of ST Households
Caste_EBC	Percentage of EBC Households
Caste_BC	Percentage of BC Households
Caste_GEN	Percentage of GENERAL Households
Hindu	Percentage of HINDU Households
Muslim	Percentage of MUSLIM Households
Buddhist	Percentage of BUDDHIST Households
Jain	Percentage of JAIN Households
NAdultMales	No. of adult males in HH
NAdultFemales	No. of adult females in HH
TotChildren_SchoolAge	No. of children in HH of school going age , 5-18 years
HH_Head_Ed~n	Education level of HH head

Table A3.3: OUTCOME VARIABLES measured at HH LEVEL at PRE-INTERVENTION VALUES

Tot_Earner_past	No. of individuals engaged in any livelihood activity at some point during 2007
Tot_PermEarner_past	No. of individuals engaged in any livelihood activity all year during 2007
Roof_Past_pucca	Percentage of HHs with Pucca Roof in 2007
Floor_Past_pucca	Percentage of HHs with Pucca Floor in 2007
Wall_Past_pucca	Percentage of HHs with Pucca Wall in 2007
Field_Toilet_past	Percentage of HHs using FIELDS for toilet in 2007
PvtToilet_past	Percentage of HHs using PRIVATE FACILITIES for toilet in 2007
PubToilet_past	Percentage of HHs using PUBLIC FACILITIES for toilet in 2007
rooms_past	No. of rooms in 2007
Land_2007	Land holding in Kathas in 2007
Chair_2007	No. of chairs in 2007
Cot_2007	No. of cots in 2007
Mobile_2007	No. of mobiles in 2007
Watch_2007	No. of watches in 2007
Cycle_2007	No. of cycles in 2007
Cows_2007	No. of cows in 2007
Buffalo~2007	No. of buffaloes in 2007
Bullock~2007	No. of bullocks in 2007
Goats_2007	No. of goats in 2007
months_insecurity_past	No. of months in 2007 during which HH faced food shortage
Act_Beating_Past (in 2nd and 4th spec)	Percentage of respondents capable of any action in 2007 due to issue: woman being beaten
Act_PDS_Past (in 2nd and 4th spec)	Percentage of respondents capable of any action in 2007 due to issue: PDS
Act_School_Past (in 2nd and 4th spec)	Percentage of respondents capable of any action in 2007 due to issue: SCHOOL
Act_Mukhiya_Past (in 2nd and 4th spec)	Percentage of respondents capable of any action in 2007 due to issue: MUKHIYA

Table A3.4 Balance in Pre-Intervention Levels of Household Outcome Variables, across matching modules

Variable	NN, with replacement				CALIPER				RADIUS				KERNEL			
	Treated	Control	t	p>t	Treated	Control	t	p>t	Treated	Control	t	p>t	Treated	Control	t	p>t
Tot_Earner_past	1.7336	1.7542	-0.85	0.396	1.7336	1.7542	-0.85	0.396	1.7336	1.7542	-0.85	0.396	1.7336	1.7542	-0.85	0.396
	1.7361	1.7278	0.34	0.732	1.7255	1.7337	-0.32	0.749	1.7255	1.7196	0.23	0.818	1.7361	1.7191	0.71	0.480
Tot_PermEarner_past	1.6342	1.6995	-2.60	0.009	1.6342	1.6995	-2.60	0.009	1.6342	1.6995	-2.60	0.009	1.6342	1.6995	-2.60	0.009
	1.6376	1.6294	0.32	0.746	1.654	1.6663	-0.46	0.643	1.654	1.6479	0.23	0.820	1.6376	1.6242	0.53	0.594
Roof_Past_pucca	.26999	.24272	1.96	0.051	.26999	.24272	1.96	0.051	.26999	.24272	1.96	0.051	.26999	.24272	1.96	0.051
	.26959	.29175	-1.54	0.125	.25748	.26745	-0.66	0.508	.25748	.25208	0.36	0.718	.26959	.26928	0.02	0.983
Floor_Past_pucca	.07081	.04241	3.85	0.000	.07081	.04241	3.85	0.000	.07081	.04241	3.85	0.000	.07081	.04241	3.85	0.000
	.0701	.06907	0.13	0.900	.05455	.05044	0.54	0.591	.05455	.04973	0.63	0.528	.0701	.07012	-0.00	0.998
Wall_Past_pucca	.27509	.26367	0.81	0.420	.27509	.26367	0.81	0.420	.27509	.26367	0.81	0.420	.27509	.26367	0.81	0.420
	.27423	.27629	-0.14	0.886	.26334	.2563	0.47	0.639	.26334	.2667	-0.22	0.824	.27423	.27411	0.01	0.994
Field_Toilet_past	.91238	.93051	-2.11	0.035	.91238	.93051	-2.11	0.035	.91238	.93051	-2.11	0.035	.91238	.93051	-2.11	0.035
	.91443	.91134	0.34	0.733	.92727	.92199	0.58	0.559	.92727	.91947	0.86	0.392	.91443	.91377	0.07	0.941
PvtToilet_past	.04381	.03321	1.72	0.085	.04381	.03321	1.72	0.085	.04381	.03321	1.72	0.085	.04381	.03321	1.72	0.085
	.04278	.03918	0.57	0.571	.0346	.03578	-0.19	0.853	.0346	.04176	-1.09	0.276	.04278	.04184	0.15	0.884
rooms_past	1.8467	1.8186	0.85	0.393	1.8467	1.8186	0.85	0.393	1.8467	1.8186	0.85	0.393	1.8467	1.8186	0.85	0.393
	1.8464	1.8459	0.02	0.988	1.8217	1.8393	-0.50	0.614	1.8217	1.8111	0.31	0.758	1.8464	1.8297	0.51	0.611
Land_2007	10.097	10.561	-0.44	0.660	10.097	10.561	-0.44	0.660	10.097	10.561	-0.44	0.660	10.097	10.561	-0.44	0.660
	10.113	12.32	-2.12	0.034	9.2373	11.262	-2.08	0.038	9.2373	10.135	-1.03	0.302	10.113	11.325	-1.21	0.227
Chair_2007	.60316	.51865	2.70	0.007	.60316	.51865	2.70	0.007	.60316	.51865	2.70	0.007	.60316	.51865	2.70	0.007
	.5933	.55619	1.18	0.238	.5607	.53842	0.67	0.501	.5607	.56975	-0.27	0.788	.5933	.58934	0.12	0.901
Cot_2007	1.7718	1.8007	-0.72	0.474	1.7718	1.8007	-0.72	0.474	1.7718	1.8007	-0.72	0.474	1.7718	1.8007	-0.72	0.474
	1.7747	1.7397	0.86	0.388	1.7548	1.7196	0.83	0.409	1.7548	1.7476	0.17	0.865	1.7747	1.776	-0.03	0.975
Mobile_2007	.35507	.25652	6.10	0.000	.35507	.25652	6.10	0.000	.35507	.25652	6.10	0.000	.35507	.25652	6.10	0.000
	.34639	.3634	-1.06	0.290	.31613	.32727	-0.68	0.499	.31613	.32655	-0.63	0.531	.34639	.35963	-0.81	0.416
Watch_2007	.32094	.23761	5.16	0.000	.32094	.23761	5.16	0.000	.32094	.23761	5.16	0.000	.32094	.23761	5.16	0.000
	.31753	.31959	-0.12	0.905	.28211	.27155	0.62	0.534	.28211	.28469	-0.15	0.881	.31753	.32113	-0.21	0.837

Variable	NN, with replacement			CALIPER			RADIUS			KERNEL		
	Treated	Control	t p>t	Treated	Control	t p>t	Treated	Control	t p>t	Treated	Control	t p>t
Cycle_2007	.36882	.30301	4.11 0.000	.36882	.30301	4.11 0.000	.36882	.30301	4.11 0.000	.36882	.30301	4.11 0.000
	.36495	.36546	-0.03 0.975	.34076	.34721	-0.38 0.702	.34076	.33244	0.49 0.622	.36495	.36299	0.12 0.905
Cows_2007	.30056	.28615	0.72 0.473	.30056	.28615	0.72 0.473	.30056	.28615	0.72 0.473	.30056	.28615	0.72 0.473
	.29897	.30876	-0.48 0.631	.28328	.31789	-1.60 0.109	.28328	.28266	0.03 0.976	.29897	.28511	0.70 0.483
Buffalo~2007	.24452	.23352	0.59 0.554	.24452	.23352	0.59 0.554	.24452	.23352	0.59 0.554	.24452	.23352	0.59 0.554
	.2433	.28454	-2.09 0.037	.24223	.28856	-2.17 0.030	.24223	.27327	-1.50 0.133	.2433	.25959	-0.87 0.386
Bullock~2007	.20122	.23301	-1.70 0.090	.20122	.23301	-1.70 0.090	.20122	.23301	-1.70 0.090	.20122	.23301	-1.70 0.090
	.20258	.22577	-1.26 0.207	.20645	.23109	-1.25 0.211	.20645	.21003	-0.19 0.853	.20258	.20787	-0.29 0.772
Goats_2007	.56139	.56975	-0.22 0.823	.56139	.56975	-0.22 0.823	.56139	.56975	-0.22 0.823	.56139	.56975	-0.22 0.823
	.56392	.56546	-0.04 0.968	.57302	.59765	-0.57 0.567	.57302	.5744	-0.03 0.974	.56392	.5612	0.07 0.945
months_insecurity_past	1.1885	1.3388	-2.51 0.012	1.1885	1.3388	-2.51 0.012	1.1885	1.3388	-2.51 0.012	1.1885	1.3388	-2.51 0.012
	1.1979	1.1876	0.17 0.863	1.2188	1.2897	-1.10 0.273	1.2188	1.2325	-0.21 0.830	1.1979	1.1314	1.14 0.255
PubToilet_past	.04381	.03628	1.20 0.229	.04381	.03628	1.20 0.229	.04381	.03628	1.20 0.229	.04381	.03628	1.20 0.229
	.04278	.04948	-0.99 0.320	.03812	.04223	-0.61 0.542	.03812	.03876	-0.10 0.923	.04278	.04439	-0.25 0.806

Table A4 ATTs, across NN, Radius and Kernel Algorithms: All households reporting women's modules considered

Variable	Nearest Neighbor, with Replacement				Radius Matching, Caliper=0.0005				Kernel Matching, Bandwidth=0.0302			
	Controls	Difference	T-stat		Controls	Difference	T-stat		Controls	Difference	T-stat	
land_owned	10.85	-0.77	-0.63		10.65	-1.28	-1.18		10.95	-0.87	-0.86	
Del_Land	0.02	-0.04	-0.55		-0.02	0.02	0.47		-0.06	0.04	1.01	
leased_in_2010	4.64	0.98	1.57		4.68	0.67	1.24		4.92	0.70	1.42	
Del_LeaseIn	0.36	-0.46	-1.87	**	0.15	-0.30	-1.17		0.18	-0.27	-1.21	
chair_current	0.74	0.01	0.29		0.70	-0.02	-0.38		0.74	0.01	0.23	
cot_current	2.05	-0.09	-1.58		2.00	-0.05	-1		2.01	-0.05	-1.05	
mobile_current	0.53	0.03	1.23		0.47	0.06	2.57	***	0.52	0.04	2.14	**
watch_current	0.41	0.00	-0.18		0.36	0.01	0.6		0.39	0.02	0.78	
cycle_current	0.48	-0.03	-1.27		0.44	-0.02	-0.74		0.44	0.00	0.14	
cows_current	0.34	0.01	0.29		0.34	-0.01	-0.29		0.32	0.03	1.1	
buffaloes_curr~t	0.27	0.02	0.59		0.30	-0.02	-0.89		0.30	-0.02	-0.76	
bullocks_current	0.20	0.03	1.02		0.22	0.02	0.74		0.22	0.01	0.29	
goats_current	0.54	0.04	0.72		0.59	0.01	0.29		0.57	0.00	0.13	
Del_Chair	0.16	0.01	0.41		0.17	-0.02	-1.1		0.15	0.02	0.98	
Del_Cot	0.20	-0.01	-0.42		0.22	-0.04	-1.61		0.22	-0.03	-1.42	
Del_Mobile	0.17	0.04	2.21	**	0.16	0.05	3.29	***	0.16	0.06	3.91	***
Del_Cycle	0.09	0.00	-0.16		0.08	0.00	-0.19		0.08	0.01	0.73	
Del_Watch	0.07	0.02	1.27		0.07	0.02	1.74	**	0.07	0.02	2.39	***
Del_Cows	0.03	0.02	1.38		0.03	0.01	0.6		0.03	0.02	1.95	**
Del_Buffaloes	0.04	0.00	0.2		0.05	-0.01	-1.26		0.05	-0.01	-1.24	
Del_Bullocks	0.01	0.01	0.92		0.01	0.01	1.45		0.01	0.01	1.88	**
Del_Goats	0.02	-0.01	-0.32		0.03	-0.02	-0.65		0.01	0.01	0.39	
Roof_Pres_Pucca	0.28	0.02	1.16		0.28	0.01	0.29		0.30	0.01	0.58	
Del_Roof	0.03	0.00	0.19		0.03	0.01	1.01		0.03	0.01	1.43	
Floor_Pres_Pucca	0.07	0.00	-0.09		0.06	-0.01	-0.89		0.08	-0.01	-0.92	
Del_Floor	0.01	-0.01	-2	**	0.01	-0.01	-2.01	**	0.01	-0.01	-1.29	
Wall_Pres_Pucca	0.32	0.00	-0.17		0.31	-0.01	-0.54		0.31	0.00	0.04	
Del_Wall	0.04	-0.01	-0.58		0.04	0.00	-0.04		0.04	0.00	0.06	

Variable	Controls	Difference	T-stat		Controls	Difference	T-stat		Controls	Difference	T-stat	
Field_Toilet_~nt	0.90	-0.01	-0.86		0.91	-0.01	-0.64		0.90	-0.01	-0.91	
Del_Field	-0.02	-0.01	-1.72	**	-0.02	-0.01	-1.88	**	-0.01	-0.02	-2.68	***
Common_Toilet~nt	0.05	0.01	0.73		0.05	0.00	0.44		0.05	0.00	0.33	
Del_Common	0.01	0.01	1.63		0.01	0.01	1.44		0.01	0.01	1.52	
Private_Toile~nt	0.05	0.00	0.46		0.04	0.00	0.45		0.05	0.01	0.94	
Del_Private	0.01	0.00	0.64		0.01	0.00	1.1		0.01	0.01	2.17	***
rooms_present	1.89	0.01	0.13		1.90	-0.02	-0.6		1.88	0.02	0.53	
food_secure_pr~t	0.32	0.03	1.17		0.35	-0.01	-0.27		0.34	0.00	0.11	
Del_Insecure	-0.02	-0.01	-0.82		-0.01	-0.02	-2.16	**	-0.01	-0.02	-2.15	**
months_insecu~nt	0.80	0.08	1.19		0.91	-0.02	-0.33		0.89	-0.01	-0.2	
Del_Month	-0.26	-0.05	-1.13		-0.24	-0.07	-1.59		-0.24	-0.06	-1.73	**
Del_Intensity1	0.68	-0.01	-0.57		0.67	-0.01	-0.33		0.68	-0.01	-0.46	
Del_Intensity2	0.14	-0.01	-0.81		0.15	-0.02	-1.31		0.14	-0.01	-1.04	
Del_Intensity3	0.18	0.03	1.39		0.18	0.02	1.52		0.19	0.02	1.44	
Tot_Earner_Pre~t	1.71	0.03	0.88		1.74	0.00	0.06		1.72	0.02	0.75	
Del_Earner	0.01	0.00	-0.73		0.01	0.00	-0.42		0.01	0.00	-0.38	
Tot_PermEarne~nt	1.62	0.03	0.84		1.67	0.00	-0.02		1.63	0.02	0.66	
Del_Perm_Earner	0.01	0.00	-0.16		0.01	0.00	0.17		0.01	0.00	0.21	
HH_Net_In_Kharif_Ag_Own	-0.01	0.01	1.09		0.00	0.00	-0.04		0.00	0.00	-0.71	
HH_Net_In_Rabi_Ag_Own	0.00	0.00	0.85		0.00	0.00	-0.07		0.00	0.00	-0.38	
HH_Net_In_Zaid_Ag_Own	0.00	-0.01	-1.32		0.00	-0.01	-1.95	**	0.00	-0.01	-1.41	
HH_Net_In_Kharif_Ag_Rent	0.00	0.00	-0.69		0.00	0.00	-1.38		0.00	0.00	-1.17	
HH_Net_In_Rabi_Ag_Rent	0.00	0.00	-1.01		0.00	-0.01	-1.56		0.00	0.00	-1.48	
HH_Net_In_Zaid_Ag_Rent	0.00	0.00	-0.38		0.00	0.00	-0.79		0.00	0.00	-1.58	
HH_Net_In_Kharif_Ag_Lab	-0.01	0.01	1.34		-0.01	0.00	0.82		-0.01	0.01	1.55	
HH_Net_In_Rabi_Ag_Lab	-0.01	0.01	1.42		-0.01	0.00	0.9		-0.01	0.01	1.69	**
HH_Net_In_Zaid_Ag_Lab	0.00	0.00	-1.05		0.00	0.00	-0.44		0.00	0.00	-0.11	
HH_Net_In_Kharif_A_Husb	0.00	0.00	0.93		0.00	0.00	1.02		-0.01	0.00	2.11	**
HH_Net_In_Rabi_A_Husb	-0.01	0.01	1.69	**	-0.01	0.00	1.78	**	-0.01	0.01	2.13	**
HH_Net_In_Zaid_A_Husb	0.00	0.00	1.61		0.00	0.00	1.75	**	0.00	0.00	1.73	**
HH_Net_In_Kharif_Nfarm	0.00	0.00	0.18		0.00	0.01	1.92	**	0.00	0.00	-1.21	
HH_Net_In_Rabi_Nfarm	0.00	0.00	0.35		0.00	0.01	2.1	**	0.00	0.00	-0.98	
HH_Net_In_Zaid_Nfarm	-0.01	0.01	2.07	**	-0.01	0.01	2.77	***	0.00	0.00	0.84	

Variable	Controls	Difference	T-stat		Controls	Difference	T-stat		Controls	Difference	T-stat	
HH_Net_In_Kharif_Lab_Urban	0.00	0.01	1.29		0.00	0.00	1.14		0.00	0.01	1.77	**
HH_Net_In_Rabi_Lab_Urban	0.00	0.01	1.08		0.00	0.00	0.65		0.00	0.00	1.25	
HH_Net_In_Zaid_Lab_Urban	0.00	0.00	0.12		0.00	0.00	-0.04		0.00	0.00	0.15	
HH_Net_In_Kharif_Lab_Rural	0.00	-0.01	-1.31		0.00	0.00	-0.04		0.00	0.00	-0.65	
HH_Net_In_Rabi_Lab_Rural	0.00	-0.01	-1.2		-0.01	0.00	0.34		0.00	0.00	-0.24	
HH_Net_In_Zaid_Lab_Rural	-0.01	0.01	1.26		-0.01	0.01	1.35		0.00	0.00	0.81	
HH_Net_In_Kharif_Sal	0.01	-0.01	-1.69	**	0.01	-0.01	-2.48	***	0.01	0.00	-1.47	
HH_Net_In_Rabi_Sal	0.01	-0.01	-1.54		0.01	-0.01	-2.43	***	0.01	0.00	-1.49	
HH_Net_In_Zaid_Sal	0.00	0.00	-0.82		0.01	-0.01	-2.33	***	0.00	0.00	-0.27	
Tot_Boy_in_Sch~l	1.06	0.01	0.22		0.99	0.06	1.52		1.04	0.04	1.09	
Tot_Girl_in_Sc~l	0.90	0.02	0.34		0.84	0.05	1.24		0.87	0.04	1.12	
savings_2010	0.24	0.71	39.59	***	0.22	0.73	51.74	***	0.24	0.71	57.01	***
Del_Saving	0.13	0.58	32.73	***	0.12	0.59	37.51	***	0.13	0.58	40.76	***
NFormal_present	0.01	-0.01	-2.71	***	0.01	-0.01	-3.39	***	0.01	-0.01	-2.82	***
Del_NFormal	0.00	-0.01	-2.5	***	0.00	-0.01	-2.88	***	0.00	-0.01	-2.67	***
Formal_present	0.09	-0.06	-5.24	***	0.08	-0.06	-6.47	***	0.09	-0.06	-7.82	***
Del_Formal	0.03	-0.02	-2.8	***	0.02	-0.02	-2.83	***	0.02	-0.02	-2.79	***
SHG_present	0.10	0.82	59.76	***	0.10	0.82	70.54	***	0.11	0.81	77.47	***
Del_SHG	0.09	0.63	38.9	***	0.08	0.63	43.4	***	0.09	0.62	46.94	***
IndebtedOldHC	0.05	-0.02	-1.88	**	0.06	-0.02	-2.65	***	0.06	-0.02	-2.72	***
HC_OldLoan_Total	1346.79	129.43	0.16		1207.54	355.87	0.46		1319.10	157.11	0.23	
IndebtedNewHC	0.44	-0.19	-8.49	***	0.43	-0.18	-9.84	***	0.44	-0.19	-11.16	***
HC_NewLoan_Total	7749.35	-3525.65	-2.38	***	8073.77	-3955.76	-3.73	***	7927.87	-3704.17	-3.98	***
N_Loan	0.69	0.48	13.07	***	0.69	0.48	15.34	***	0.69	0.49	17.17	***
Total_Borrowed	11031.02	813.00	0.47		11077.16	690.71	0.51		11041.88	802.14	0.68	
Tot_Loan_Health	0.29	0.15	5.45	***	0.28	0.17	7.08	***	0.27	0.17	7.73	***
Tot_Amt_Health	3965.45	529.23	0.34		3910.79	804.09	0.71		3449.09	1045.58	1.07	
Tot_Loan_Marital	0.15	-0.01	-0.38		0.15	-0.01	-0.42		0.15	-0.01	-0.58	
Tot_Amt_Marital	2897.10	-782.93	-1.54		2742.44	-664.76	-1.6		2834.64	-720.48	-1.94	**
Tot_Loan_Food	0.07	0.04	2.85	***	0.07	0.04	2.87	***	0.07	0.04	3.33	***
Tot_Amt_Food	544.22	48.00	0.31		833.34	-269.18	-0.83		693.56	-101.34	-0.35	
Tot_Loan_House	0.06	0.05	3.82	***	0.06	0.03	3.04	***	0.06	0.04	4.3	***
Tot_Amt_House	1048.61	127.86	0.47		1059.81	-40.54	-0.17		1277.04	-100.56	-0.45	

Variable	Controls	Difference	T-stat		Controls	Difference	T-stat		Controls	Difference	T-stat	
Tot_Loan_Durab~s	0.03	0.04	3.37	***	0.03	0.03	3.6	***	0.04	0.03	3.68	***
Tot_Amt_Durables	539.73	-52.69	-0.29		613.84	-119.20	-0.63		560.66	-73.63	-0.44	
Tot_Loan_School	0.01	0.01	0.76		0.01	0.00	0.78		0.01	0.00	0.13	
Tot_Amt_School	48.61	83.85	1.16		56.15	91.83	1.57		76.28	56.18	1.1	
Tot_Loan_Debt	0.01	0.08	9.96	***	0.01	0.08	10.61	***	0.01	0.08	11.38	***
Tot_Amt_Debt	52.85	673.44	7.16	***	43.63	708.99	7.22	***	46.56	679.73	7.85	***
Tot_Loan_Lives~k	0.01	0.06	7.53	***	0.01	0.06	7.67	***	0.01	0.05	7.97	***
Tot_Amt_Livest~k	132.25	426.42	3.68	***	174.14	394.61	4.07	***	165.63	393.05	4.52	***
Tot_Loan_Shop	0.02	0.04	3.86	***	0.02	0.03	3.88	***	0.02	0.04	5.43	***
Tot_Amt_Shop	380.29	228.86	1.26		318.74	231.75	1.68		302.59	306.56	2.46	***
Tot_Loan_Machine	0.00	0.01	1.77	**	0.00	0.01	1.78	**	0.01	0.01	1.77	**
Tot_Amt_Machine	538.42	-328.49	-0.69		485.11	-273.77	-0.95		628.94	-419.01	-1.67	**
Self_Response	0.73	0.17	8.8	***	0.73	0.17	10.72	***	0.74	0.16	11.65	***
age_marriage	15.88	-0.08	-0.29		15.76	-0.08	-0.33		15.89	-0.09	-0.4	
son_education	9.34	0.50	2.26	**	9.22	0.46	2.43	***	9.37	0.47	2.71	***
daughter_educa~n	7.94	0.42	1.96	**	7.94	0.22	1.2		8.10	0.25	1.52	
signature_lite~e	0.35	0.33	13.33	***	0.34	0.33	16.41	***	0.35	0.33	18.79	***
sign_literate	0.19	0.03	1.84	**	0.17	0.04	2.37	**	0.18	0.04	3.16	**
job_card	0.24	0.00	-0.08		0.25	0.00	0.16		0.24	0.00	-0.23	
kirana_present	0.83	0.04	2.09	**	0.83	0.04	2.73	***	0.84	0.03	2.28	**
Del_Kirana	0.05	0.01	1.23		0.06	0.01	0.83		0.06	0.01	1.33	
pds_present	0.65	0.00	0.21		0.66	-0.01	-0.36		0.65	0.00	0.06	
Del_PDS	0.04	0.02	1.55		0.04	0.02	2.22	**	0.04	0.02	2.42	***
health_present	0.81	0.06	3.53	***	0.82	0.05	3.09	***	0.82	0.05	4.09	***
Del_Health	0.06	0.03	2.02	**	0.06	0.02	1.8	**	0.06	0.03	2.58	***
neighbor_present	0.95	0.02	2.55	***	0.96	0.02	2.39	***	0.96	0.01	2	**
Del_Neighbor	0.04	0.02	2.09	**	0.05	0.01	1.43		0.06	0.01	1.08	
relative_present	0.85	0.05	3.19	***	0.85	0.04	2.93	***	0.86	0.03	2.78	***
Del_Relative	0.04	0.03	2.76	***	0.05	0.03	2.49	***	0.05	0.02	2.07	**
panchayat_pres~t	0.03	0.05	5.84	***	0.03	0.06	6.68	***	0.04	0.05	6.47	***
Del_Panchayat	0.00	0.02	3.92	***	0.00	0.02	3.69	***	0.00	0.02	4.05	***
decision_cook~nt	0.93	0.00	-0.4		0.93	0.00	0.08		0.93	0.00	0.02	
Del_Cook	0.04	-0.01	-0.69		0.04	0.00	-0.24		0.03	0.01	0.91	

Variable	Controls	Difference	T-stat		Controls	Difference	T-stat		Controls	Difference	T-stat	
decision_dura~nt	0.43	0.08	3.56	***	0.41	0.09	4.48	***	0.44	0.07	4.18	***
Del_Durable	0.03	0.01	0.68		0.03	0.01	0.94		0.02	0.02	2.25	**
decision_pers~nt	0.92	0.01	0.99		0.93	0.00	0.45		0.93	0.01	0.64	
Del_Personal	0.05	-0.01	-0.93		0.05	-0.01	-0.98		0.05	0.00	-0.27	
decision_heal~nt	0.70	0.04	2.09	**	0.70	0.03	1.55		0.71	0.04	2.16	**
Del_Health	0.06	0.03	2.02	**	0.06	0.02	1.8	**	0.06	0.03	2.58	***
decision_educ~nt	0.70	0.10	4.54	***	0.69	0.09	4.83	***	0.71	0.09	5.33	***
Del_Education	0.06	0.01	1.27		0.06	0.02	1.6		0.06	0.02	2.26	**
dec_live_pres	0.50	0.05	2.25	**	0.50	0.06	2.78	***	0.49	0.05	3.04	***
Del_Livelihood	0.02	0.02	2.59	***	0.02	0.02	2.37	***	0.02	0.02	2.98	***
dec_employ_pres	0.53	0.06	2.78	***	0.52	0.08	3.81	***	0.52	0.08	4.21	***
Del_Employment	0.03	0.02	1.85	**	0.02	0.01	2.07	**	0.02	0.02	3.06	**
decision_loan~nt	0.58	0.23	10.32	***	0.58	0.23	12.58	***	0.58	0.23	14.22	***
Del_Loan	0.03	0.07	7.09	***	0.03	0.06	6.45	***	0.02	0.07	8.37	***
decision_poli~nt	0.20	0.09	4.61	***	0.19	0.10	6.01	***	0.21	0.09	5.73	***
Del_Politics	0.01	0.04	5.78	***	0.01	0.03	5.24	***	0.01	0.04	5.66	***
beating_actio~nt	0.72	0.11	5.43	***	0.74	0.09	5.67	***	0.75	0.08	5.63	***
Del_Beat_Act	0.16	0.06	3.61	***	0.15	0.08	5.22	***	0.15	0.07	5.3	***
pds_action_pre~t	0.38	0.05	2.33	**	0.39	0.03	1.62		0.38	0.05	3.06	***
Del_PDS_Act	0.10	0.05	3.45	***	0.10	0.05	3.76	***	0.09	0.06	4.75	***
school_action~nt	0.35	0.06	2.73	***	0.34	0.05	2.62	***	0.36	0.06	3.26	***
Del_School_Act	0.09	0.06	4.17	***	0.10	0.05	4.19	***	0.10	0.06	4.86	***
mukhiya_actio~nt	0.26	0.03	1.52		0.26	0.02	1.35		0.26	0.03	1.78	**
Del_Mukhiya_Act	0.06	0.04	3.16	***	0.06	0.03	2.45	***	0.06	0.03	3.31	***
ASelfBPres	0.36	0.11	4.68	***	0.36	0.09	4.52	***	0.38	0.08	4.58	***
Del_Self_beat	0.08	0.07	4.22	***	0.08	0.05	4.01	***	0.09	0.05	4.25	***
AWomBPres	0.39	0.16	6.93	***	0.40	0.14	6.88	***	0.40	0.15	8.72	***
Del_Wom_beat	0.10	0.12	7.22	***	0.10	0.12	8.3	***	0.09	0.13	10.14	***
ASelfPDSPres	0.18	0.01	0.66		0.17	0.02	1.05		0.18	0.02	1.14	
Del_Self_pds	0.05	0.01	0.95		0.05	0.01	0.98		0.05	0.02	1.93	**
AWomPDSPres	0.17	0.05	3.08	***	0.15	0.07	4.52	***	0.14	0.08	6.09	***
Del_Wom_pds	0.05	0.06	5.27	***	0.04	0.07	6.18	***	0.04	0.07	7.27	***
ASelfSPres	0.16	0.02	1.49		0.14	0.03	2.2	**	0.15	0.03	2.2	**

Variable	Controls	Difference	T-stat		Controls	Difference	T-stat		Controls	Difference	T-stat	
Del_Self_school	0.05	0.03	2.34	**	0.05	0.03	2.83	***	0.05	0.03	2.9	***
AWomSPres	0.17	0.06	3.66	***	0.15	0.06	4.26	***	0.16	0.07	5.47	***
Del_Wom_school	0.05	0.07	5.46	***	0.05	0.06	5.38	***	0.05	0.06	6.32	***
ASelfMPres	0.10	0.01	0.72		0.09	0.01	0.87		0.09	0.02	2.01	**
Del_Self_mukhiya	0.03	0.01	1.56		0.03	0.01	1.13		0.03	0.01	2.06	**
AWomMPres	0.12	0.04	2.92	***	0.11	0.04	3.45	***	0.12	0.05	3.99	***
Del_Wom_mukhiya	0.03	0.05	5.15	***	0.03	0.04	4.72	***	0.03	0.04	5.7	***

Significance at 95% denoted by "", significance at 99% denoted by "***"**

References

Agarwal, B. (1997). "Bargaining"and Gender Relations: Within and Beyond the Household. *Feminist economics*, *3*(1), 1-51.

Anderson, S., and Eswaran, M. (2009). What determines female autonomy? Evidence from Bangladesh. *Journal of Development Economics*, *90*(2), 179-191.

Banerjee, A. V., Duflo, E., Glennerster, R., and Kinnan, C. (2009). The Miracle of Microfinance?: Evidence from a Randomized Evaluation.

Bobonis, G. J. (2009). Is the allocation of resources within the household efficient? New evidence from a randomized experiment. *Journal of Political Economy*, *117*(3), 453-503.

Caliendo, M., and Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of economic surveys*, *22*(1), 31-72.

Dehejia, R. H., and Wahba, S. (2002). Propensity score-matching methods for nonexperimental causal studies. *Review of Economics and statistics*, *84*(1), 151-161.

Dehejia, R. (2005). Practical propensity score matching: a reply to Smith and Todd. *Journal of Econometrics*, *125*(1), 355-364.

Feigenberg, B., Field, E. M., and Pande, R. (2010). *Building social capital through microfinance* (No. w16018). National Bureau of Economic Research.

Heckman, J. J., Ichimura, H., and Todd, P. (1998). Matching as an econometric evaluation estimator. *The Review of Economic Studies*, *65*(2), 261-294.

Heinrich, C., Maffioli, A., and Vazquez, G. (2010). A primer for applying propensity-score matching. *SPD Working Papers*

Kabeer, N. (2002). Resources, agency, achievements: Reflections on the measurement of women's empowerment. *Development and change*, *30*(3), 435-464.

Kazianga, H., and Udry, C. (2006). Consumption smoothing? Livestock, insurance and drought in rural Burkina Faso. *Journal of Development Economics*, *79*(2), 413-446.

Lancaster, G., Maitra, P., and Ray, R. (2006). Endogenous Intra-household Balance of Power and its Impact on Expenditure Patterns: Evidence from India. *Economica*, *73*(291), 435-460

Luke, N., and Munshi, K. (2007). Social affiliation and the demand for health services: Caste and child health in South India. *Journal of development economics*, *83*(2), 256-279.

Luke, N., and Munshi, K. (2011). Women as agents of change: Female income and mobility in India. *Journal of Development Economics*, *94*(1), 1-17.

Pitt, M. M., and Khandker, S. R. (1998). The impact of group-based credit programs on poor households in Bangladesh: does the gender of participants matter?. *Journal of political economy*, *106*(5), 958-996.

Rosenbaum, P. R., and Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41-55.

Smith, J. A., and Todd, P. E. (2001). Reconciling conflicting evidence on the performance of propensity-score matching methods. *The American Economic Review*, *91*(2), 112-118.

Thomas, D., Contreras, D., and Frankenberg, E. (2002). Distribution of power within the household and child health. *Santa Monica: RAND. Mimeo*.