ELSEVIER
www.elsevier.com/locate/worlddev

World Development Vol. 66, pp. 443–456, 2015
0305-750X/

http://dx.doi.org/10.1016/j.worlddev.2014.08.024

Evaluating Impacts of Watershed Development Program on Agricultural Productivity, Income, and Livelihood in Bhalki Watershed of Bardhaman District, West Bengal

NIRUPAM DATTA*

Indira Gandhi Institute of Development Research (IGIDR), Mumbai, India
International Food Policy Research Institute (IFPRI), New Delhi, India

Summary. — Watershed development facilitates in reducing the vulnerability of farm income to weather-induced shocks in rain-fed lands in India. The present paper estimates homogeneous as well as non-homogeneous effects of watershed development on farmers demonstrating a huge discrepancy in estimated values between two. It shows that initial differences in resource endowments, access to formal credit, education level, and caste membership can result in inequalities in the impact of watershed development programs on targeted parameters. The analysis also establishes the sensitivity of the estimates to the general level of multivariate imbalance as well as univariate imbalance in confounding covariates.

Key words — India, West Bengal, watershed, impact evaluation, coarsened exact matching (CEM)

1. INTRODUCTION

Dry-land agricultural system, covering around three-fifths of 142 million hectares of cultivable land provides about 45% of India's food requirements (GOI, 2012). This agricultural system being basically rain-fed, suffers from a variety of problems. High variability in the amount of rainfall received in short intensive spells separated by long dry periods during the monsoons exposes the standing crops to the risks of both water stress and floods, resulting in low yields if not outright crop failures (Rockstrom *et al.*, 2010). Agricultural lands under rain-fed areas also suffer from poor soil quality, having low organic and mineral content due to leaching and erosion of top soil that is carried along with surface run-off before being dumped into downstream water storage structures leading to siltation as well as poor quality of collected water (Garg, Wani, Barron, Karlberg, & Rockstrom, 2012). These ecological problems are further exacerbated by high population pressure, abject poverty, and almost non-existent or poorly maintained infrastructure such as metalled roads, water supply, irrigation, education, health services, credit facilities, extension services, *etc* (Bouma & Scott, 2006).

In the years following independence when the need of ensuring food security for the country was very much pressing, the policy makers decided to concentrate the scarce resources to usher in a green revolution, in areas that were relatively well endowed with better water and soil resources, quite inadvertently ignoring the rain-fed areas. Though this approach led to an increase in production of food-grains and agricultural income it also led to an increase in inequality between the areas of intervention and rain-fed areas. In recent times, emerging indications of stagnant if not decelerating growth rates in the areas where the green revolution first took place has forced the policy makers to rethink and turn their attention toward the development of rain-fed agriculture to raise yield levels as well as production of food grains (Garg *et al.*, 2012). Fan, Hazell, and Haque (2000) have shown that as the rates on returns to increasing levels of investment tends to decrease in irrigated areas, it actually increases in the rain-fed areas.

The effort that started with technical centric approach on soil conservation in the 1970 and water conservation in the 1980s veered off toward a participatory approach of all the stakeholders in the project areas in the 1990s and beyond (Joshi, Jha, Wani, Sreedevi, & Shaheen, 2008). The present participatory approach incorporates all the elements of soil and water conservation along with improvement in agricultural and social infrastructure, market and credit access as well as introduction of new agricultural technologies all bundled together under the generic name of watershed development. Hydrologically, watershed is a catchment area from where the water after rainfall, flows through a network of streams, gullies and then discharges through a common outlet into a particular drainage system such as a river, ranging from a few hectares to several thousands of hectares of surface area. Watershed development refers to "conservation, regeneration, judicious and harmonious use of human and natural (like land, water, plants, animals) resources within a particular watershed through implementation of 'ridge-to-valley' approach involving all the stakeholders, in a sustainable manner to generate a stream of various services such as food, fodder, fuel wood, water for agriculture and domestic purposes,

* The author would like to express gratitude to Mr. Fakir Bagdi, who had worked tirelessly with the author, taking him across the villages, introducing him to the villagers besides providing housekeeping services during the author's stay in the survey area. The author would also like to express gratitude to the Bhalki Jalabibhajika Unnayan Committee (BJUC), NABARD officials for their logistical support. Finally, the author is grateful to the thesis committee members, Dr. Vijay Laxmi Pandey, Dr. Srijit Mishra, and Dr. Rabindranath Bhattacharya, Prof. Gary King, Harvard University and two anonymous referees for their helpful comments and suggestions for making the work more perfect.This research is supported in part by a Grant from the Indira Gandhi Institute of Development Research, Mumbai. Final revision accepted: August 19, 2014.

reducing soil erosion, recharging groundwater table *etc*." (NABARD, 2006). As the boundaries of a single macro-watershed do not align with those of basic administrative and geo-political units most of the time, it is divided into several micro-watersheds to make them overlap as much as possible with these boundaries. This also makes the task of managing the development program much easier as it facilitates easier resolution of conflict of interests among various stakeholders. Development of watersheds has been believed to be instrumental in raising yields, especially that of dry-land crops, crop diversification, extension of cropping activities during the dry season, and employment opportunities in rain-fed areas, where resource degradation is a serious problem (Hope, 2007; Kerr, 2002). Sharma, Rao, Vittal, Ramakrishna, and Amarasinghe (2010) using the water balance method found that the amount of usable rainwater in 225 dominant rain-fed districts is more than sufficient for providing one round of supplemental irrigation over an area of more than 20 million hectares even during drought years which if managed properly can lead to increase in yield of up to 50% for dry-land crops such as pulses and oilseeds. Sahrawat, Wani, Pathak, and Rego (2010) in an experimental project, found that soil and water conservation measures with proper nutrient management not only helped to increase productivity but also enhanced soil quality.

However, all the benefits claimed in favor of watershed development do not go uncontested. Bouma and Scott (2006) analyzing household surveys and focus groups discussion from four watershed projects in peninsular India found that watershed development does not necessarily lead to an increase in yields of dry land crops but leads to a shift toward more water-intensive crops due to existing imperfect markets and lack of proper developed varieties of dry-land crops and access to affordable credit for making profitable investments to improve agricultural activities. Glendenning and Vervoort (2010, 2011) used the water balance method to estimate the per cent of harvested rain water in rain water-harvesting structures (RWHs) that was recharged into the groundwater table in the Arvari river basin area of Rajasthan. They found that there was a huge gap in the potential recharge and the actual recharge values. Only around 7% of harvested rain water was recharged into the groundwater table with the rest probably being accounted for by soil storage and lateral flow into surrounding areas. They also found contrary to the claims, increasing the number of RWHs beyond a certain point will have no additional benefit in terms of groundwater recharge and will drastically reduce the runoff into downstream areas. At the optimal level, when the level of sustainability is highest, increase in irrigated area will actually reduce the level of resilience. However, there is no doubt that in times of drought RWHs somewhat acts as a sort of buffer against crop failures. Bouma, Biggs, and Bouwer (2011) while studying the impact of watershed development in the Musi sub-basin area on the downstream reservoirs that supply water to Hyderabad city found that while watershed development has led to an increase in cropping activities with diversification toward vegetable crops leading to an increase in income of the farmers, it has led to a drastic fall in runoff into the water reservoirs, more so during dry years that may adversely impact the supply of water to Hyderabad city. The net returns to watershed development activities after accounting for the fall in surface runoff downstream actually turn out to be negative, underscoring the need to factor in the uncertainties that loom large over such developmental activities. But Garg *et al*. (2012) argued that while development in the Musi sub-basin may have reduced the flow into the water reservoirs downstream it has also reduced the level of siltation and has also led to an improvement in the quality of water. In fact the greatest benefits from the project are available when there is only in situ soil and water conservation interventions rather than having no intervention. Similarly, in the Bundelkhand region, Singh, Garg, Wani, Tewari, and Dhyani (2014) found that though integrated watershed development project reduces storm flow it increased the base flow thereby reducing the adverse impacts of watershed project on downstream areas in terms of runoff. Similarly in the Rajasamadhiyala watershed project in Gujarat, Sreedevi *et al*. (2006) found that watershed development not only did lead to an increase in level of groundwater in the project area but also in the control watersheds several kilometers downstream apart from increase in yields, cropping intensity as well as crop diversity.

Estimates for watershed development in a particular year vary from half a billion dollars as reported in Bouma and Scott (2006) to around two billion dollars in Joshi *et al*. (2008). Given the importance of watershed treatment as an important tool in accelerating the development of rain-fed regions of the country, it becomes essential to assess the impact of such projects (Hope, 2007).

2. THEORY OF CHANGE

To begin with, Figure 1 below provides a schematic diagram, depicting the causal pathways through which watershed development affects cropping activities. In the first stage of development, treatment measures in terms of construction of new and renovation of existing water-storage structures such as contour trenches on the ridges, reservoirs, inter-connected pits in the fields and checks dams at the lower end of the drainage lines, planting of trees in the upper parts of the watershed and lining of drainage channels with grass are undertaken (Garg, Karlberg, Barron, Wani, & Rockstrom, 2011; Singh *et al*., 2014; Sreedevi *et al*., 2006). These reduce surface run-off that occurs during short but intensive phases of rainfall during the monsoons. Reduction in surface run-off triggers the change leading to a reduction in washing away of the top soil along with an increase in retention of water in soil and water bodies, also due to a reduction in siltation, another effect of reduction in surface run-off, and triggering further changes through three interrelated but different pathways at the third stage. While both reduction in soil erosion and increased water retention lead to increased moisture content in the soil, more percolation of water downward due to increased water retention leads to an increase in the groundwater level thereby increasing its harnessing potential for use during the dry season for both agricultural and domestic purposes (Rockstrom *et al*., 2010). On the other hand, reduction in soil erosion alone leads to improvement in soil quality through increased activities of soil micro-organisms. Agricultural inputs especially chemicals as well as HYV seeds are more responsive to the availability of adequate moisture and good soil quality; improvements in soil characteristics and water availability lead to increased demand for these inputs. As the increased response of these inputs will lead to more productivity, it will lead to more demand for paid labor for increasing cropping activities (Sreedevi *et al*., 2006). Also as HYVs are more susceptible to pest attacks, it also leads to an increased demand for pesticides. Moreover as water becomes also available during the dry season, there is also a trend of intensification of cropping activities by farmers. The possibility of a shift from mono-cropping to multiple cropping practices especially toward cultivation of multiple vegetable

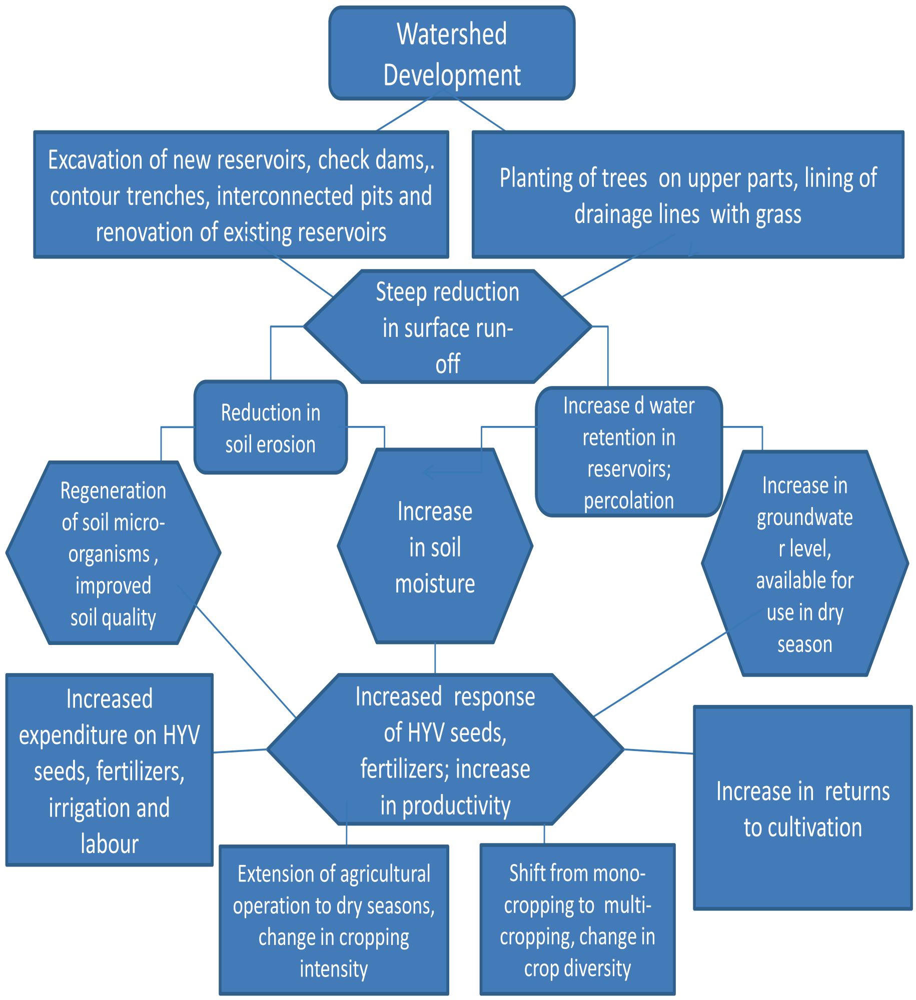

Figure 1. *Pathways through which watershed development impacts agriculture.*

crops that require assured water supply also rises substantially (Bouma *et al.*, 2011). In all, there will be a general increase in costs of cultivation owing to either a combination of or all the factors explained above. To capture the impacts of watershed development on activities that preclude the employment of a large number of specialists equipped with sophisticated instruments to measure hydrological and various environmental and ecological parameters, it is quite justified to estimate the changes in terms of well-defined and easily measurable indicators such as cropping intensity, crop diversity, agricultural income, and costs of cultivation.

3. METHODOLOGICAL ISSUES – WAY FORWARD

To assess the impact properly, it is necessary to measure the indicators both in the presence as well in the absence of the treatment by collecting data from both the treated and control micro-watersheds within the same macro watershed. This is important when the amount of resources allocated in a treatment micro-watershed is quite significant. It may be possible that the changes that may be potentially ascribable to watershed development may have occurred due to some fundamental changes in the project areas that have nothing to do with the project, say a change in climatic conditions, better market situation, or government policies leading to opening up of better opportunities. In that case weak or lack of proper comparison framework may lead to overestimation of the effects of watershed development leading to biased policy conclusions and hence wrongful scaling up of such interventions on a larger scale (Clemens & Demombynes, 2011). Though the literature on watershed impact assessment in India is quite large and varied, most of them suffer from lack of comparison frameworks that are required for proper evaluation of impacts (Kerr, Pangare, Pangare, & George, 2000). Most of these studies rely only on the data collected from the treated watershed only, for assessing the impact which at times can be grossly misleading (Clemens & Demombynes, 2011). According to the author's knowledge, two studies that have attempted to analyze the impact of watershed development using well-established quantitative techniques are Kerr *et al.* (2000) and Hope (2007). Kerr *et al.* (2000) uses an instrumental variable approach to determine the factors responsible for selection of the villages under different watershed development programs in Maharashtra and Andhra Pradesh. The paper then uses the predicted values of participation of villages under different projects to determine the impact on different indicators such as net returns to cultivation in the program villages, extent of erosion in the drainage lines, investments in land improvement, *etc*. In this study most of the indicators as well the treatment status are analyzed at the village level. Hope (2007) used household data from both the treated as well as the control micro-watersheds in Madhya Pradesh and matched the treated households with the control ones using the predicted probabilities obtained from the logistic regression, regressing access to land and threshold time taken to collect drinking water on several socio-economic variables, as estimated propensity scores. After matching the treated with control households, impact of watershed development on farm income in the Rabi and Kharif seasons and time taken to collect drinking water were found. Though these papers have facilitated the understanding of the impacts of watershed development to a significant extent, they did not capture effects on other outcomes such as costs of cultivation, total revenue, cropping intensity, and crop diversity that are very important from the view point of the planners who implement these projects as these indicators also reflect the intensity of resource use as well as diversification of crop portfolio that is important in reducing weather-induced crop-based income shocks. These studies also suffer from the drawbacks of dependence on modeling assumptions being made when the true model through which the data have been generated is unknown (Ho, Imai, King, & Stuart, 2007). They also estimated the impacts of watershed development assuming a homogeneous treatment effect i.e., equal effect on the treated units irrespective of their characteristics, which is too simplistic an assumption. The present study tries to assess how modeling assumptions and level of imbalance between the treatment and control units can impact the results obtained. It also makes an attempt to assess non-homogeneous treatment effects i.e., unequal effects of treatment depending on the characteristics of the treatment units. The following sections discuss in brief about the study area, information gathered through open-ended discussions with the project officers, members of the villages watershed committee as well as select villagers about the impacts of the project at the village level followed by information on data collected for the study and the methodology to analyze the same. The final section discusses the results of the analysis with policy implications.

4. BRIEF DESCRIPTION OF THE STUDY AREA

The area of study, the Bhalki region lying at 23°27′28″N 87°37′28″E and its immediate surrounding areas, are located in Ausgram-II block in the western part of the Burdwan district, predominantly a backward area with semi-arid climate and red lateritic sandy soils. Summers are hot and winters are cool. Average annual rainfall is around 1,200 mm most of which is received during the monsoons (personal communication with block officials, 2011–12). Around two-thirds of the total population in the area belong to the backward community. The project village, Bhalki as well as the control villages are located on the upper parts of the drainage basin of the Ajay-Kunnur river system, tributaries of the Damodar river, at a distance of 45 kilometrrs from district headquarters, Burdwan, 25 kilometers from Durgapur, an important industrial township and 20 kilometers from Bolpur in the Birbhum district. Table 1 gives a break-up of the total area in Bhalki under different land uses.

Table 1. *Geographical characteristics of the Bhalki micro-watershed*

Characteristics	Hectares
Total area	954.15
Forest area	359.74
Wasteland	214.15
Cultivable area	316.27

Watershed development measures have been undertaken in the village during the period 2002–08, utilizing the funds from NABARD's micro-watershed development program. The area is a bit undulated with a slope of less than 3%. Most of the farm households are marginal land holders, many of them working as share croppers. The treatment micro-watershed contains forested lands locally known as *Jungle Mahal*, covering around a third of the total micro-watershed area. Before the commencement of the micro-watershed development project, agricultural activity in the treatment village was limited to the rainy season only. Mainly, paddy was grown having a yield of 2–2.5 tonnes per hectare.

5. PROJECT DEVELOPMENT

In early 2001, some villagers along with the local block development officer (BDO) came forward to organize the inhabitants of Bhalki into different SHGs to undertake micro-watershed development program. Through meetings, awareness was generated by the change agents among the inhabitants about the importance of soil and water conservation measures and the benefits that can be shared by them. People were convinced; many of them joined as members of self-help groups (SHGs). During the pre-project phase, saplings of *arjun*, *akashmoni*, *sonajhuri*, *shirish*, bamboo were planted on the wastelands covering around 100 hectares. The work was done basically by the members of the SHGs, many of whom were women from the backward communities. NABARD officials, impressed with the progress of work released around INR 270,000 for payment toward the completion of job in confidence build-up phase. However, the members of the SHGs refused the money and instead demanded it to be used for setting up a mobile microfinance institution under the auspices of the watershed committee. It was created to cater basically to the working capital needs of the small farmers. In all, NABARD sanctioned INR 5.5 million for watershed development. Table 2 lists information on the work carried out under watershed development.

During the final implementation phase of the project that lasted for around six years, more soil and water conservation structures as well as income-generating avenues were created. As the slope of the land in Bhalki is less than 3%, focus was laid much on the excavation of large reservoirs and inter-connected system of ponds, laid out in a way such that when the drainage lines are carrying excess run-off, a lateral outlet would force the water inside these ponds. Apart from these, other conservation structures such as contour trenches and contour bunds were also created in the afforested areas as well as on the cultivated lands. To prevent erosion of bunds, trees such as Burma teak, sal, *shirish*, *arjun*, guava were planted that also provided a long-term asset-based source of income to the members of SHGs that maintain these structures. Pisciculture has been promoted in the area; training has been provided to

Table 2. *Project work in the Bhalki micro-watershed*

Project work	Numbers/area covered in hectares
Afforestation of wastelands	200 Ha
Water reservoirs	2 Nos.(Large)
Ponds	16 Nos.
Pisciculture	2 Ponds (Large)
Orchards	10 Nos.
Lacquer project	1 No.
Check dam	1 No.

members of different self-help groups. Currently, three to four large ponds that have been excavated under watershed development are being utilized for pisciculture apart from providing irrigation. Orchards of mango, cashew, guava, and jackfruit covering several hectares of land have also been developed. One of these orchards, developed on an abandoned garbage dump was nurtured carefully with pitcher-based irrigation, plant by plant. In these orchards, during the gestation period, vegetables and sunflower are being grown for sustaining the SHG members looking after them. Besides these, social afforestation and nursery development have also been taken up on a massive scale.

6. IMPACT OF MICRO-WATERSHED DEVELOPMENT ON THE VILLAGERS

Micro-watershed development program has brought enormous changes in the lives of the households in Bhalki. The social forestry project has created a huge asset base to the tune of at least INR 25–30 million for the SHG members. In the first phase, 10 SHGs have earned INR 4 million from the sale of trees. Previously, owners of lands in the upper parts did not bother about using their lands for any productive use. Success of the afforestation project has encouraged them to seek saplings from the nurseries run by the SHGs for planting by themselves.

Development and renovation of water bodies for water storage as well as percolation into the ground as well as conserving of soil moisture have brought a huge change in the agricultural potential in the area. Cultivation of different vegetable crops have been taken up that are sold in the nearby three wholesale markets. Development of fisheries, providing means of livelihood to around 20 families have led to an annual turnover of INR 1.5–2 million from each pond. The orchards are being looked after properly. They are expected to yield on a commercial scale within the next two years.

Due to sandy soil, the rate of percolation from the water bodies is quite high leading to drying up within 2–3 months after the withdrawal of monsoons. Though there are no permanent arrangements for monitoring, personal communication with watershed and NABARD officials has revealed a rise in the level of the water table from 300 feet to 100–120 feet in some places in the watershed and 180 feet on an average.

Though the farmers of Bhalki have been able to flood the local markets with their produce quite successfully, expansion of that opportunity to bigger markets still remains to be achieved.

7. DATA

For the purpose of the present study, a census was conducted of all the farm households in the treatment and the contiguous nine control villages during November–December 2011 after the end of the Kharif season and during April 2012 toward the end of the Rabi season. The census has information on 226 agricultural households from the treatment micro-watersheds and 439 agricultural households from the control ones, thereby totaling 665 respondents. Information was collected on socio-economic variables such as caste, religion, education status of the household members, employment patterns, net area under cultivation, access to formal source of credit for cropping activities, participation in rural employment guarantee scheme, NREGA, and common farm household assets. Agricultural information collected included those on the crops cultivated in Kharif and Rabi seasons and the area under each crop, yield, information on marketing of produce, costs under different input heads including seeds, fertilizers, pesticides, traction, irrigation, and different categories of paid labor required for cultivation on a crop-by-crop basis. As the project did not involve any sort of randomized control design and as baseline data on various socio-economic characteristics of the households from the treated and control micro-watersheds are not available, proper care needs to be taken in choice of an appropriate framework that controls for the effect of the potential confounders on the indicators of interest apart from choice of appropriate pre-treatment variables i.e., variables who values are not impacted as a result of the treatment.

8. METHODOLOGY FOR ESTIMATION OF TREATMENT EFFECT

For estimating the impact, the treatment is assumed to be ignorable conditioned on observed confounders and that every treated unit receives the same treatment. A fixed causal effect is a function of potential outcome defined as:

$$Y_i(1) - Y_i(0)$$

It is the difference in the potential outcomes under treatment and control for unit i that are not necessarily observed (Ho *et al.*, 2007). To estimate the causal effect due to treatment, we require for the treated unit under question, the value of the potential outcome both under the treatment as well as under control status for evaluation. But the unit under question can only be in either of these two groups at any point of time. So for every unit under study, one of these potential outcomes will always be unobserved, known as the fundamental problem of causal inference (Holland, 1986).

In conventional observational studies, quantitative assessment of impacts of treatment effect usually requires modeling assumptions and specifications. However, there is no well-established method to deduce the correct functional form (Ho *et al.*, 2007). The reasoning is equally applicable for matching methods that use estimated propensity score as a matching variable whereby the true propensity score-generating model is unknown and hence there is no benchmark against which to compare the estimated models.

To reduce the dependence of the estimands of interest on modeling assumptions and empirical specifications and hence to obtain less researcher-induced bias and more reliable estimates, the present study first pre-processes the raw survey data through a reduction in imbalance in covariates among the treated and control units by balancing of the empirical multivariate distribution of the covariates defined by the following metric (Iacus, King, and Porro, 2011a):

$$L_1(f, g; H) = \frac{1}{2} \sum_{l_1 l_2 \ldots\ldots\ldots\ldots l_k \varepsilon H(X)} |f_{l_1 l_2 \ldots\ldots l_k} - g_{l_1 l_2 \ldots\ldots l_k}|$$

where

L_1 = multivariate imbalance measure

$H(X) = \prod_{i=1}^{k} H(X_i)$, multidimensional histogram constructed from the set of cells generated by the Cartesian product of $H(X_i)$ s

$H(X_i)$ = sets of intervals into which the supports of the variables X_i s have been cut or coarsened (the length of which is less than or equal to the range of the values for X_i – it is usually user defined depending on the contexts). This is the maximum level of imbalance set ex ante for matching.

f and g are the empirical multivariate frequency distributions for the treated and control units respectively and $f_{l_1 l_2 \ldots l_k}$ and $g_{l_1 l_2 \ldots l_k}$ are the relative multivariate frequencies for observations belonging to the cells with coordinates $l_1, l_2, \ldots\ldots\ldots\ldots l_k$ as confounding pre-treatment variables for the treated and control units, respectively.

Compared to other matching methods, the method mentioned above not only reduces imbalance in means of the covariates between and treated and control units but also imbalances in higher moments of the empirical distributions and other non-linearities and interactions due to better overlapping. The remaining imbalances within the matched strata in the values between the matched treated and control units are then controlled through parametric modeling, albeit, with reduced model dependence. Theoretically, pre-matching as well as post-matching the value of L_1 lies in the closed interval [0, 1]. Prior to matching, if $L_1 = 0$ then the dataset is perfectly balanced. If $L_1 = 1$, the dataset is totally imbalanced and no sort of matching exercise can make any difference. Hence the imbalance measure itself provides an answer whether it will be worthwhile to undertake the exercise of matching.

This method, known as coarsened exact matching (CEM) method relies on the coarsening of the values of covariates into different well-defined intervals based on substantive knowledge of the problem, for better overlapping between the treated and control groups. It satisfies the monotonic imbalance bounding principle as discussed in Iacus, King, & Porro, 2011b whereby the level of imbalance chosen ex ante for one variable will not alter the maximum level of imbalance chosen for the other variables thus reducing the uncertainty regarding the level of imbalance on other variables which is not available in any other prevalent matching methods. It also meets congruence principle between the data and analysis spaces.

Coarsened exact matching also eliminates the pre-matching requirement of many matching methods of restricting the data set to the common empirical support for both the treated and control groups as for the observations within a coarsened stratum containing treated and control units, no extrapolation beyond the data is involved. It is also invariant to measurement error for the variables provided the maximum error is less than or equal to the maximum width of the coarsened intervals for the particular variable and respects the resulting strata boundaries. By choosing the level of imbalance ex ante the researcher bounds the degree of model dependence as well as estimation error for the treatment effect. In the absence of perfect experimental designs, it is essential that the number of units in the control group should be at least as much as the number of units from the treatment group. As CEM like any other matching method, involves pruning off of incompatible units from both the treated and control groups it is essential to have more number of control units at least in the ratio of around 1:2.

9. INDICATORS OF INTEREST AND ESTIMANDS

For the current study, the indicators of interest whose changes are to be estimated because of the treatment on the matched sub-sample of treated and control units are:

a) Δ Annual net returns = (Net returns from agriculture in treated watershed area) – (Net returns from agriculture in control watershed)
b) Δ Annual revenue = (Revenue from sale of crops in treated watershed) – (Revenue from sale of crops in control watershed)
c) Δ Annual costs of cultivation = (Costs of Cultivation in treated watershed) – (Costs of Cultivation in control watershed). All measured in terms of per hectare of cultivated land
d) Δ Cropping intensity = (Gross Cropped Area/Net Cropped Area)$_{\text{treated watershed}}$ – (Gross Cropped Area/Net Cropped Area)$_{\text{control watershed}}$
e) Δ Cropping Diversity $= \left\{1-\sum_{i=1}^{n}(C_i/\sum C_i)^2\right\}_{\text{treated watershed}}$

$-\left\{1-\sum_{i=1}^{n}(C_i/\sum C_i)^2\right\}_{\text{control watershed}}$

where $\sum_{i=1}^{n}(C_i/\sum C_i)^2$ = Hirschman–Herfindahl Index
C_i = income from crop i

The estimands for the causal effects on indicators can be defined as:

$$SATT = \frac{1}{n_T}\sum_{i \varepsilon T} TE_i$$

where $TE_i = Y_i(1) - \widehat{Y}_i(0)|X_i$
n_T = total number of treated units in the original sample.

The above estimand is applicable only when all the treated units are matched. As in our case, when all treated units are not matched then $SATT$ changes to $LSATT$ or local sample average treatment for the treated and the estimands consequently changes to:

$$LSATT = \frac{1}{m_T}\sum_{i \varepsilon T^m} TE_i$$

where m_T = number of matched treated units
T^m = subset of matched treated units.

For the unmatched treated units, we extrapolate via some model, the estimands on the matched treated and control units to obtain virtual control units and then obtain an estimate of the estimands for the unmatched treated units. The overall $SATT$ is then calculated as the weighted average of the two estimates for matched and unmatched treated units.

10. RESULTS

(a) *Summary statistics of raw data*

Tables 3 and 4 show the summary statistics for interval scale and categorical pre-treatment variables that are potential confounders, respectively. On an average, the number of household members engaged in agriculture varies from a little over 2 in the case of control micro-watersheds to a little over 2.5 in the case of treated micro-watersheds. Household members engaged in agricultural activities have average highest education of 5.89 years in treated micro-watersheds compared to 7.42 years in control micro-watersheds. Average highest farming experience of household members currently engaged in agriculture varies very little among the treated and control micro-watersheds, being around 28 years. Farm households on an average cultivate 0.91 hectare and 1.04 hectares of land in treated and control micro-watersheds, respectively and hence are mostly marginal farmers as per land size classification used by the state. In terms of formal sources of credit, lower proportion of farm households in treated micro-watersheds have access compared to those of control areas. This observation in some way raises questions about the claims of the watershed officials in the program villages regarding easy availability of cheap formal micro-credit through various self-help groups. In fact, almost every agricultural household

Table 3. *Summary statistics for interval variables*

Variable	Unit of measurement	Treatment status	Number of observations	Mean value
Farm hands	No. of household members	Treated	226	2.52(1.40)
		Control	439	2.15(1.35)
Highest educated farm hand	Years of education	Treated	226	5.89(4.65)
		Control	439	7.42(4.56)
Experience of oldest farm hand	Years of farming experience	Treated	226	27.88(14.36)
		Control	439	27.92(14.12)
Land	Area cultivated by farm household	Treated	226	0.91(1.00)
		Control	439	

Figures in parentheses denote standard deviation.

Table 4. *Summary statistics for categorical variables*

Variable	Treatment status	No. of observations	Proportion of households
Access to formal credit	Treated	226	18.58
	Control	439	24.37
Belonging to backward class	Treated	226	80.97
	Control	439	69.70

surveyed in the treated micro-watershed, feigned ignorance when asked even about the existence of such groups. Also, majority of households belong to the backward community in treated than in control micro-watersheds. As the comparison between the households of the treated and control watersheds show that differences do exist in the values of agriculturally important parameters, there is a pressing need to match the households in the treated watershed as closely as possible with similar households from the control watersheds.

(b) *Matching results*

For the purpose of our analysis, we first try to reduce the imbalance in the values of six confounding covariates: land, farm hands per household, experience of the oldest farm hand, highest educated farm hand, access to formal source of credit, and whether the household belongs to the state-defined broad social category of backward community or not, between treated and control groups. Access to markets has not been included in the list of confounding covariates as the market for paddy for both the treatment and control micro-watershed households have been state-funded cooperative societies or local wholesale markets. As far as access to existing wholesale

Table 5. *Measures of imbalance for initial raw data*

Multivariate imbalance measure: L1 = 0.872
Percentage of local common support: LCS = 7.3%
Univariate imbalance measures:

Variable	Statistic	Type	L1	Min	25%	50%	75%	Max
Farm hands	−0.3696353	(diff)	0.1582035	1	0	0	0	1
Highest education	1.5230512	(diff)	0.1369968	0	5	1	1	2
Farm experience	0.0351866	(diff)	0.0255408	1	3	0	−2	−7
Area cultivated	0.8261934	(diff)	0.1049247	0	0	0.5	1	−9
Formal credit	0.0578951	(diff)	0.0578951	0	0	0	0	0
Backward	−0.1126958	(diff)	0.1126958	0	−1	0	0	0

Table 6. *Matching results for automated coarsening of covariates*

	Treatment	Control
All	226	439
Matched	98	113
Unmatched	128	326

Multivariate imbalance measure: L1 = 0.642
Percentage of local common support: LCS = 23.3%
Univariate imbalance measures:

Variable	Statistic	Type	L1	Min	25%	50%	75%	Max
Farm hands	0.00E+00	(diff)	4.86E−17	0	0	0	0	0
Highest education	−2.89E−02	(diff)	1.02E−02	0	0	0	0	0
Farm experience	4.91E−01	(diff)	1.02E−02	0	0	0	0	5
Area cultivated	3.11E−01	(diff)	6.46E−02	0	1	0.5	0	0
Formal credit	0.00E+00	(diff)	5.90E−17	0	0	0	0	0
Backward	−1.11E−16	(diff)	6.94E−17	0	0	0	0	0

markets is concerned, they are the same for the households in both the treated and control watersheds. There was no evidence from, both the data collected and also from personal communication with the farmers about any special marketing arrangements, formal or informal to have been worked out between the traders of the wholesale markets and farmers. Whatever agricultural transactions take place are spot transactions; goods are sold to only those buyers who are ready to provide the best prices on the day of transaction concerned. Similarly, due to universal participation both in terms of number of days as well as proportion of households from the treated and control watersheds, participation in NREGA has also been kept out of the list of potential confounders. Table 5 shows the initial level of multivariate imbalance for the overall raw data and for the individual covariates.

For a starting point of reference, the matching algorithm CEM developed by Iacus, King, & Porro, 2009, is allowed to automatically coarsen the values of the covariates to match the treated and the control units where the maximum tolerable level of ex-ante imbalance is fixed automatically. Table 6 shows the results of the automated matching method. There are 98 treated units matched to 113 control units with a post matching multivariate imbalance of 0.642 compared to 0.872 for the original raw data. The region of common support for covariates between matched treated and control units is 23.3% of the data space compared to just 7.3% for the raw data.

In order to get more number of matched treated as well as control units, we now coarsen the *Highest Education* variable by grouping values into different categories as per classification of different levels of education: uneducated (0 year), primary (1–4 years), middle high school (5–8 years), high school (9–12 years) and university (13 years and above). We next coarsen the *Area Cultivated* variable in two different ways, (a) first by grouping the values as per farmer-type land size classification: (i) Marginal (>0<=1), (ii) Small (>1<=2), (iii) Medium (>2<=4) and (iv) Medium-Large (>4<=10) and (b) secondly, in intervals of 1 hectare each from 0.32 hectares to 4 hectares and the last interval by 4 hectares, 4.01 hectares to 8.01 hectares. We then carry out two different matching exercises using the CEM algorithm with the same level of coarsening for the *Highest Education* variable but two different levels of coarsening for the *Area Cultivated* variable thereby yielding different levels of multivariate imbalance as well as a number of matched treated and control units. Tables 5 and 6 respectively show the outcomes of the different matching exercises.

In Table 7 corresponding to a multivariate imbalance measure of 0.699 post-matching, we obtain 164 control and 125 treated units whereas in Table 8, for a multivariate imbalance measure of 0.583, we obtain 106 treated and 119 control units. While in the Table 5, the region of common support for the covariates between matched treated and control units are 19.2% of the data space, it is 25.4% in Table 6.

Table 7. *Matching results for user-defined coarsening*

	Treated	Control
All	226	439
Matched	125	164
Unmatched	101	275
Multivariate imbalance measure: L1 = 0.699		
Percentage of local common support: LCS = 19.2%		
Univariate imbalance measures:		

Variable	Statistic	Type	L1	Min	25%	50%	75%	Max
Farm hands	4.44E−16	(diff)	0.00E+00	0	0	0	0	0
Highest education	−2.55E−02	(diff)	2.78E−17	0	0	1	0	0
Farm experience	1.54E-01	(diff)	1.33E−02	0	1	0	0	5
Area cultivated	−9.85E−02	(diff)	4.92E−02	0	0.5	0	−1.25	8
Formal credit	1.39E−17	(diff)	0.00E+00	0	0	0	0	0
Backward	1.11E−16	(diff)	1.39E−17	0	0	0	0	0

Table 8. *Matching results for user-defined coarsening*

	Treated	Control
All	226	439
Matched	106	119
Unmatched	120	320
Multivariate imbalance measure: L1 = 0.583		
Percentage of local common support: LCS = 25.4%		
Univariate imbalance measures:		

Variable	Statistic	Type	L1	Min	25%	50%	75%	Max
Farm hands	0	(diff)	3.82E−17	0	0	0	0	0
Highest education	−0.0861635	(diff)	8.50E−17	0	0	0	0	0
Farm experience	0.4356469	(diff)	1.89E−02	0	2	0	0	5
Area cultivated	−0.0223158	(diff)	3.99E−02	0	0.5	0	-0.5	8
Formal credit	0	(diff)	6.25E−17	0	0	0	0	0
Backward	0	(diff)	6.94E−17	0	0	0	0	0

From Tables 7 and 8 it is clear that the level of imbalance as measured by difference in means of values of the covariates between matched treated and control units within the matched strata for *Highest Education* as well as *Area Cultivated* is several times lower than that fixed ex ante. In fact the level of imbalance post matching has improved in the case of *Area Cultivated* variable under the user-chosen coarsening levels compared to that of automated coarsening where the level of coarsening is much less.

In the case of *Highest Education* variable the level of imbalance is lesser in the case where multivariate imbalance measure is 0.699 compared to the case of automated coarsening-induced multivariate imbalance measure of 0.642; it is however larger in the case where the multivariate imbalance measure is 0.583. But the imbalance achieved is much lower than that set ex ante. The logic applies equally to other matched covariates that are automatically coarsened (stratum intervals are not shown).

(c) *Estimation of treatment effects*

We now use the matched data obtained from different matching exercises with different levels of coarsening of covariates to arrive at the estimates of local sample average treatment effect (*LSATT*) for the treated that utilizes only the matched subsamples of treated and control units and global sample average treatment effect for the treated by including the unmatched treated units in the analysis. We estimate both homogeneous (constant treatment effect across treated households) as well as non-homogeneous (non-constant treatment effect across treated households) treatment effects i.e., random effects within the matched strata. While for estimating the homogeneous effect, linear effects model has been used which can be defined as:

$$Indicator\ of\ Interest = \alpha + \beta_i X_i + \gamma\ Treatment + \varepsilon$$

where γ gives an estimate of homogeneous effect, X_i s are the pre-treatment variables that are potential confounders, and ε is the error term.

For estimating the non-homogeneous effect, linear random effects model has been used which can be written as:

$$Indicator\ of\ Interest = \alpha + \beta_i X_i + \gamma\ Treatment + U_j + \varepsilon_j$$

where γ gives an estimate of homogeneous effect, X_i s are the pre-treatment variables that are potential confounders, ε_j are individual unit-specific error terms and U_j measures the difference in impact between the individual unit j and the average homogeneous treatment effect γ. Sum of γ and U_j then gives an estimate of the non-homogeneous treatment effect.

Compared to previous studies on the impact of watershed development including those that have a more structured analytical framework, the current study makes allowance for differential treatment effect on the treated that may arise out of, but may not be necessarily limited to individual household's socio-economic characteristics and social learning. Here social learning is assumed to get in only when changes in agro-economic structure of the treated area begin to manifest as a result of micro-watershed development e.g., introduction of vegetables forvegetable cultivation in treated areas.

Tables 9–13 show the estimates of treatment effects on unit net returns, unit sales, unit cost, cropping intensity, and crop diversity. Apart from the matched data set *Automated* where the covariate values have been automatically coarsened with

Table 9. *Estimates of impact on net returns per hectare (in INR)*

Variable	Matching method	Multivariate imbalance measure	*SATT*	Effect	Point estimate	*p*-Value
Unit Net Returns	Automated	0.642	Local	Homogeneous	4737.25	0.32
				Non-Homogeneous	4584.00	0.00
			Global	Homogeneous	7176.38	0.05
				Non-Homogeneous	9552.06	0.00
	Match I	0.583	Local	Homogeneous	7372.88	0.06
				Non-Homogeneous	6938.12	0.00
			Global	Homogeneous	8270.88	0.03
				Non-Homogeneous	9650.81	0.00
	Match II	0.699	Local	Homogeneous	5738.38	0.07
				Non-Homogeneous	5547.81	0.00
			Global	Homogeneous	8186.44	0.06
				Non-Homogeneous	9597.00	0.00

Table 10. *Estimates of impact on crop income per hectare (in INR)*

Variable	Matching method	Multivariate imbalance measure	*SATT*	Effect	Point estimate	*p*-Value
Unit Sales	Automated	0.642	Local	Homogeneous	13263.75	0.01
				Non-Homogeneous	13452.75	0.00
			Global	Homogeneous	13990.00	0.00
				Non-Homogeneous	17335.44	0.00
	Match I	0.583	Local	Homogeneous	15549.44	0.00
				Non-Homogeneous	14104.88	0.00
			Global	Homogeneous	14827.00	0.00
				Non-Homogeneous	17255.88	0.00
	Match II	0.699	Local	Homogeneous	15296.12	0.00
				Non-Homogeneous	14339.75	0.00
			Global	Homogeneous	16416.62	0.00
				Non-Homogeneous	17344.25	0.00

Table 11. *Estimates of impact on costs of cultivation per hectare (in INR)*

Variable	Matching method	Multivariate imbalance measure	*SATT*	Effect	Point estimate	*p*-Value
Unit Cost	Automated	0.642	Local	Homogeneous	8526.50	0.00
				Non-Homogeneous	8517.06	0.00
			Global	Homogeneous	6813.62	0.00
				Non-Homogeneous	7537.81	0.00
	Match I	0.583	Local	Homogeneous	8176.56	0.00
				Non-Homogeneous	8023.44	0.00
			Global	Homogeneous	6556.12	0.00
				Non-Homogeneous	7410.25	0.00
	Match II	0.699	Local	Homogeneous	9557.75	0.00
				Non-Homogeneous	9138.94	0.00
			Global	Homogeneous	8230.19	0.00
				Non-Homogeneous	7462.00	0.00

Table 12. *Estimates of impact on cropping intensity*

Variable	Matching method	Multivariate imbalance measure	*SATT*	Effect	Point estimate	*p*-Value
Cropping Intensity	Automated	0.642	Local	Homogeneous	1.61	0.66
				Non-Homogeneous	0.29	0.71
			Global	Homogeneous	−0.26	0.09
				Non-Homogeneous	−3.66	1.00
	Match I	0.583	Local	Homogeneous	2.26	0.48
				Non-Homogeneous	0.94	0.07
			Global	Homogeneous	−0.14	0.04
				Non-Homogeneous	−4.11	1.00
	Match II	0.699	Local	Homogeneous	−0.12	0.97
				Non-Homogeneous	−1.58	0.99
			Global	Homogeneous	−0.74	0.22
				Non-Homogeneous	−3.77	1.00

Table 13. *Estimates of impact on crop diversity*

Variable	Matching method	Multivariate imbalance measure	*SATT*	Effect	Point estimate	*p*-Value
Kharif Diversity	Automated	0.642	Local	Homogeneous	0.02	0.33
				Non-Homogeneous	0.02	0.00
			Global	Homogeneous	0.00	1.00
				Non-Homogeneous	0.01	0.03
	Match I	0.583	Local	Homogeneous	0.01	0.74
				Non-Homogeneous	0.00	0.43
			Global	Homogeneous	0.02	0.45
				Non-Homogeneous	0.00	0.44
	Match II	0.699	Local	Homogeneous	0.01	0.73
				Non-Homogeneous	0.00	1.00
			Global	Homogeneous	0.01	0.58
				Non-Homogeneous	0.01	0.01

a post-matching multivariate imbalance measure of 0.642, we have two other matched data sets *Match I* with a post matching imbalance measure of 0.583 and *Match II* with a post matching imbalance measure of 0.699.

Table 9 shows the local *SATT* estimates on annual unit net returns. Under the assumption of homogeneous treatment effect across matched strata, it ranges from INR 4737.25 per hectare in *Automated* to INR 7372.88 per hectare in *Match I* to INR 5738.38 per hectare in *Match II*. The global estimates for the respective data sets are INR 7176.38 per hectare, INR 8270.88 per hectare, and INR 8186.44 per hectare. Relaxing the constant treatment effect, un-weighted random effects models are estimated within each stratum and then results are averaged across the stratum with appropriate weights to arrive at the non-homogenous local *SATT* estimates that stand at INR 4584.00 per hectare for *Automated*, INR 6938.12 per hectare for *Match I,* and INR 5547.81 per hectare for *Match II*. In the case of global *SATT*, the estimates for the respective matched data sets are INR 9552.06 per hectare, INR 9650.81 per hectare, and INR 9597.00 per hectare.

Table 10 shows the local *SATT* estimates on annual unit sales. In the case of constant treatment effect across matched strata, it ranges from INR 13263.75 per hectare in *Automated* to INR 15549.44 per hectare in *Match I* to INR 15296.12 per hectare in *Match II*. The global estimates for the respective data sets are INR 13990.00 per hectare, INR 14827.00 per hectare, and INR 16416.62 per hectare. Corresponding non-homogenous local *SATT* effects are INR 13452.75 per hectare for *Automated*, INR 14104.88 per hectare for *Match I,* and INR 14339.75 per hectare for *Match II*. For global *SATT*, the estimates for the respective data sets are INR 17335.44 per hectare, INR 17255.88 per hectare, and INR 17344.25 per hectare.

Table 11 contains the local *SATT* estimates on annual unit cost. Constant treatment effect ranges from INR 8526.50 per hectare in *Automated* to INR 8176.56 per hectare in *Match I* to INR 9557.75 per hectare in *Match II*. The global estimates for the respective data sets are INR 6813.62 per hectare, INR 6556.12 per hectare, and INR 8230.19 per hectare. Corresponding non-homogenous local *SATT* effects are INR 8517.06 per hectare for *Automated*, INR 8023.44 per hectare for *Match I,* and INR 91438.94 per hectare for *Match II*. For global *SATT*, the estimates for the respective data sets are INR 7537.81 per hectare, INR 7410.25 per hectare, and INR 7462.00 per hectare.

Table 12 shows changes in cropping intensity. Assuming constant treatment effect across matched strata in the case of local *SATT* it ranges from an increase of 1.61% in *Automated* to 2.26% in *Match I* to a fall of 0.12% in *Match II*. The global estimates for the respective data sets are −0.26%, −0.14% and −0.74%. Corresponding non-homogenous local *SATT* effects are 0.29% for *Automated*, 0.94% for *Match I* and −1.58% for *Match II*. For global *SATT*, the estimates for the respective data sets are −3.66%, −4.11, and −3.77%. None of the estimates are however statistically significant.

For the effect on crop diversity as displayed in Table 13, we find that there is a minute increase in diversity ranging from 0.01 to 0.02 under both constant and non-constant treatment effect assumptions across strata both for local as well as global *SATT*. However excluding for those under global non-constant treatment effect in the case of *Automated* and *Match II* as well as local constant treatment effect in the case of *Automated*, none of the estimates is statistically different from zero.

Looking back at the estimands for all the five indicators of interest above, we can easily find out that under local *SATT*, estimates of non-homogeneous treatment effect have been mostly less than those of homogeneous treatment effect. Under constant treatment effect condition, we assume away the differential abilities of the individual households arising out of their social class affiliation, wealth base, educational attainment, access to credit as well as number of household farm hands. NABARD's micro-watershed development program is basically targeted toward the backward regions; as the area under current study is also backward agriculturally as well as socially, with more than two thirds of the households belonging to the backward community, such a simplistic assumption leads to overestimation of effects as the results show. With non-homogeneous treatment effect assumption allowing for differential response to treatment, current results brings out the inequality that is evidently arising in terms of achievements in all the five indicators .This is clear even without analysis of the sub-sets of the original data set on the basis of different social categories.

But when we turn to estimates of global *SATT* for annual unit sales, annual unit net returns as well as annual unit costs of cultivation, we find those under non-homogeneous treatment assumption are almost always greater than those under constant treatment effect assumption approach. Even the absolute difference between estimates of non-homogeneous treatment and homogeneous treatment is more in the case of global *SATT* than that under local *SATT*. In other words, the difference not only changes sign as one moves from local to global estimates it also becomes larger. Again these differences in estimates between local and global *SATT*s for both non-homogeneous as well as homogeneous treatment effect vary across matched data sets. While in the case of annual unit net returns and annual unit sales, the difference is lowest in the case of *Match I* and highest in the case of *Match II* when local *SATT* is considered; in the case of global *SATT*, they are lowest in the case of *Match I* for annual unit net returns and *Match II* for annual unit sales and highest in the case of *Automated* for both indicators.

In the case of annual unit costs of cultivation, while the estimates under non-homogeneous treatment effect are more than those under the homogeneous effect, in the case of global *SATT*, unlike annual unit net returns and annual unit sales, global estimates are lower than corresponding local estimates. Under the assumption of constant treatment effect, the difference between local and global *SATT* is lowest in the case of *Match II* and highest in the case of *Automated*. Corresponding difference under non-constant treatment effect assumption is lowest in the case of *Match I* and highest in the case of *Match II*.

For the indicators, cropping intensity as well as crop diversity, both in cases of local and global *SATT*s, estimates under non-constant treatment effect are either equal or lower than those under constant treatment effect. In the case of cropping intensity the estimates change sign from positive to negative as one moves from local to global *SATT* for matched datasets *Automated* and *Match I*. The difference between local and global *SATT* under both types of treatment effects are lowest in the case of *Match II* and highest in the case of *Match I*. In the case of crop diversity, the difference between local and global *SATT* under constant effects assumption is lowest in the case of *Match II* and highest in the case of *Automated*, it is lowest in the case of *Match I* while it is higher and same under both *Automated* and *Match II* under non-constant treatment effect.

11. SIMULATIONS

We now compare CEM to a popularly used matching method propensity score matching using the steps given in Iacus *et al.*, 2011b by setting the true homogeneous treatment effect on annual unit costs of cultivation to INR 1000 and generating values of annual unit costs through the following highly non-linear form:

$Costs = 1000^{*}T + 0.1^{*}\exp\ (0.7^{*}(Area\ \ Cultivated) + 0.7^{*}(Highest\ \ Education)) + \varepsilon$ where $\varepsilon \sim N\ (0,\ 10)$. The value of T, the treatment variable is assigned to each observation on the basis of the true propensity score e, given by:

$e_i = \text{logit}^{-1}\ \ \{1 + 0.1^{*}\grave{v} + 0.01^{*}(Area\ \ Cultivated) - 0.01^{*}(Farm\ Experience) + 0.01^{*}(Farm\ Hand) + 0.3^{*}(Education\_Highest)\}$

where $\grave{v}$ is the linear predictor of a mis-specified logistic model for estimating propensity score given by:

$\grave{v} = \alpha + \beta\ \ (Area\ \ Cultivated) + \gamma\ \ (Farm\ \ Experience) + \delta\ (Backward) + \varepsilon\ (Credit) + \zeta\ (Farm\ \ Hand) + \eta\ (Highest\ \ Education)$

In each of the 200 replications of the data generation process used in Diamond and Sekhon (2013), each observation i is assigned a treatment according to $T_i \sim$ Bernoulli (e_i) thereby allowing the number of pre-match treated and control units to vary across replications. *SATT* estimators are then compared based on the difference in means (RAW), the nearest neighbor propensity score matching (PSC) and CEM using automated coarsening criterion. Comparison is based on the following parameters: bias (BIAS), standard deviation (SD), root mean square errors (RMSE) that are averaged over 200 Monte Carlo replications along with average number of

Table 14. *Comparison of bias, standard deviation, root mean square error, measure of imbalance L_1 for the raw data (RAW), nearest neighbor propensity score matching (PSC) and CEM, and number of matched treated and control units with values averaged over 200 Monte Carlo replications*

	Bias	SD	RMSE	Treated	Control	L_1
RAW	$-1.62^{*}10^{13}$	$6.9^{*}10^{13}$	$7.07^{*}10^{13}$	215	450	0.85
PSC	$-3.65^{*}10^{13}$	$1.17^{*}10^{13}$	$1.22^{*}10^{13}$	215	203	0.83
CEM	$6.86^{*}10^{2}$	27381.12	27321.2	78	77	0.73

matched treated and control units for PSC and CEM. Table 14 shows the results of the simulation exercise.

From the table it is very clear that CEM outperforms nearest neighborhood propensity score matching several times over in terms of all the parameters. Compared to propensity score matching, CEM is found to reduce the overall multivariate imbalance further in the raw data. From the table, it is evident that in the case of propensity score matching, the number of treated units matched is indifferent from that of the original raw data while it is substantially lower in the case of CEM. From the insights obtained from the analysis and as discussed above, the perils of retaining all the treated units intact including those for which there are not valid counterfactuals, are well understood. The simulation exercises also confirm these.

12. DISCUSSION

The present study, while assessing the impact of micro-watershed project has clearly shown that unquestioned dependence on modeling assumptions for estimating the impacts of treatment from totally as well as partially imbalanced observational data can adversely affect informing of the policy-making process. Varying levels of multivariate as well as univariate imbalance between the treated and control household units will have varying degrees of discrepancy between values of local and global *SATT* across different modeling assumptions regarding treatment effect. Hence sticking to a specific modeling assumption can bias decision making which is a matter of serious concern when the question of allocation of scarce resources to different types of development projects is involved.

In our analysis, we find that the sensitivities of the estimates of the five indicators of interest tend to vary across the three matched data sets in altogether different ways for both types of treatment effects. In the case of annual unit net returns and annual unit sales for constant effect, we find that the difference is most in the case of *Match II* which has the highest level of multivariate imbalance post matching followed by *Automated* and then *Match I*, in line with the descending overall level of multivariate imbalance. This is an example of general sensitivity of different modeling specifications to different levels of multivariate imbalance. On the other hand, in the case of annual unit costs of cultivation, the difference is most in the case of *Automated* followed by *Match I* and *Match II*. We find that the difference actually goes down with increasing level of multivariate imbalance which is quite perplexing. But going a little deeper we find that for the confounding covariates *Farm Hands*, *Highest Education*, *Formal Credit* and *Backward*, the level of imbalance post matching is highest for *Automated* followed by those in *Match I* and *Match II*; the difference in the estimates also follow the same pattern. Thus this case shows sensitivity of the modeling assumptions to covariate specific level of imbalance and not overall multivariate imbalance. Again for cropping intensity and crop diversity under constant effect assumption, we find that that the estimates actually turn negative from local to global *SATT*. In this case, however, the difference is most in the case of *Match I* followed by *Automated* and *Match II*. We find that though *Match I* has the lowest level of multivariate imbalance, it has the highest level of imbalance for the covariates *Farm Experience*, *Formal Credit* and *Backward* – again a case of covariate-specific sensitivity.

When we turn to the estimates of non-constant treatment effects, the situation is more complex. In the case of annual unit net returns the difference is most in the case of *Automated* followed by *Match II* & *Match I*; the pattern of difference seems to follow the level of imbalance in *Area Cultivated* across the three data sets. In the case of annual unit sales, the difference is again most in the case of *Automated* but followed by *Match I* and *Match II*. Here the pattern of difference seems to follow the pattern of difference in covariates *Farm Hands*, *Highest Education* and *Backward*. In the case of annual unit costs, the difference is most in the case of *Match II* followed by *Automated* and *Match I*, following the usual pattern of overall multivariate imbalance. In the case of cropping intensity, the pattern of difference is similar to those under constant treatment effect. Thus for the same data set, the pattern of difference in local and global *SATT* estimates differ between constant and non-constant treatment specifications.

We also find that the estimates of non-constant treatment effects are lower than their constant counterparts for all the indicators excepting for cropping intensity for local *SATT* than those under global *SATT*. This points out once again to the serious challenges the policy makers have to encounter in the case of observational studies such as this when they have to make a value judgement such as whether watershed development has led to any change in inequality of outcomes between broad social groups. There will be instances as demonstrated here when they will be crippled to arrive at any unanimous decision. Modeling specifications using all the treated and control units as in the case of pure parametric modeling or retaining all the treated units and pruning off the control units as in the case of nearest neighbor propensity score matching, make the verdict ambiguous. As it is not always possible to maintain perfect randomized control designs in the case of watershed development due to political and ethical considerations it is very important to clearly specify the assumptions about the data-generating process while drawing policy conclusions from the same. As per the author's knowledge no other earlier studies at least in the Indian context have voiced any concern regarding the sensitivities of their results either to the scope of study, data collected, and the modeling assumptions. In a way, the present study poses a question to the results that have been obtained in previous studies and hence serves as an important caveat in that regard.

Looking back at the discussion at the beginning of the study regarding the various benefits of watershed development on the livelihoods of the people, we find from the current study that some of them may not hold true in all the cases due to interplay of several local factors. While improvements have been accomplished in terms of economic parameters such as unit net returns to cultivation, unit sales for the farm households in the treated micro-watershed, improvements in terms of ecologically as well as sustainability wise critical agronomic parameters – cropping intensity and crop diversity, have been marginal at the best if not negative (though statistically being insignificant) compared to what is being usually claimed in favor of watershed development. An intuitive explanation can be given for such lackluster results based on the circumstances that prevailed in the area during the period of study as well as the pre-existing economic conditions. Though not reported here, treatment effect on unit cost for the Kharif season shows that there has been a substantial increase in expenditure in treated area over that of control areas that almost entirely accounts for overall increase in expenditure for the treated households in the entire season. As paddy is grown mainly during the Kharif season in these areas, this change in terms of expenditure can be interpreted as the response of the households in the treated areas to pull in all the resources for paddy cultivation that requires much less effort compared to other crops such as vegetables and are also less susceptible to pests and diseases. During the season under study, the wholesale market for paddy has been down due to a bumper harvest as well as large inventory accumulated from the preceding two seasons; farmers were unable to sell their harvest. After government intervention, they were able to sell their produce to the farmers' cooperatives but the payment was delayed by at least a few months. Combining this fact with poor access to cheap formal sources of credit, reflecting failures in terms of community efforts in terms of mobilizing resources for providing adequate credit at affordable rates and increased expenditure made by the treated households than the non-treated ones may have discouraged the farmers in the treated areas to undertake further expenditure for Rabi cultivation. Discussions with farmers from the treated areas somewhat point toward this.

Reinforcing this dampening effect is policy-induced switching off of a major causal pathway to affect cropping activities through the blanket ban on exploitation of groundwater resources since 2006 that was lifted only during 2012 monsoons. We have discussed earlier in the article that a major change arising out of watershed development is a steep increase in groundwater levels in the treated micro-watershed. But this potential increase in groundwater level could not be harnessed because of lack of public investments in bore as well as tube-wells which are capital intensive and hence prohibitively expensive for resource constrained farmers of this area. Though some gray private investments have been made in some fields and these are where all vegetable cultivation takes place, they fall much below the potential that can be sustainably exploited. Barron, Noel, and Mikhail (2009) while reviewing various case studies of watershed development found that successful examples of intervention consisted of at least a few if not more number of effective measures rather than just one or two. The present study reinforces the finding.

Such a prevailing situation despite adoption of soil and water conservation measures is indeed a matter of concern as inability to diversify as well as intensify the cropping activities due to resource constraints and ill-conceived public policies can prove to be more costly from risk and vulnerability point of view, undermining its very basis. It also raises a serious question regarding equitable distribution of resources such as groundwater as only few are able to exploit the resources successfully at the expense of others. This has, in a way, to do with the faulty water resource regulating institutional mechanisms at vogue in Indian states that lead to over exploitation of groundwater leading to its depletion. As the actual resource cost of water is hardly reflected in the user charges there is always an adverse incentive to indulge in wasteful irrigation technologies as well as shifting to water-intensive crops (Shiferaw, Reddy, & Wani, 2008). It can also severely cripple the ability of the system to be resilient during dry periods or dry years when supplemental irrigation will prove to be more crucial in saving the crops as well as ensuring food security.

AUTHOR NOTE

Nirupam Datta is a Ph.D. candidate at the Indira Gandhi Institute of Development Research, Mumbai and is now working as Research Analyst at the International Food Policy Research Institute (IFPRI) South Asia Regional Office, New Delhi. Views expressed in the paper do not necessarily reflect those of the management of either IGIDR or IFPRI and the author takes full responsibility of the claims being made in it.

REFERENCES

Barron, J., Noel, S., & Mikhail, M. (2009). *Review of agricultural water intervention impacts at the watershed scale: A synthesis using the sustainable livelihoods framework*. Stockholm Environment Institute.

Bouma, J. A., & Scott, C. (2006). The possibilities for dryland crop yield improvement in India's semi-arid regions: Observations from the field. *Comprehensive Assessment Discussion Paper No. 3*. Colombo, Sri Lanka: Comprehensive Assessment Secretariat.

Bouma, J. A., Biggs, T. W., & Bouwer, L. M. (2011). The downstream externalities of harvesting rainwater in semi-arid watersheds: An Indian case study. *Agricultural Water Management, 98*, 1162–1170.

Clemens, M. A., & Demombynes, G. (2011). When does rigorous impact evaluation make a difference? The case of the Millennium Villages. *Journal of Development Effectiveness, 3*(3), 305–339.

Diamond, A., & Sekhon, J. S. (2013). Genetic matching for estimating causal effects: A general multivariate matching method for achieving balance in observational studies. *The Review of Economics and Statistics, 95*(3), 932–945.

Fan, S., Hazell, P., & Haque, T. (2000). Targeting public investments by agro-ecological zone to achieve growth and poverty alleviation goals in rural India. *Food Policy, 25*, 411–428.

Garg, K. K., Karlberg, L., Barron, J., Wani, S. P., & Rockstrom, J. (2011). Assessing the impact of agricultural water interventions at the Kothapally watershed, Southern India. *Hydrological Processes, 26*(3), 387–404.

Garg, K. K., Wani, S. P., Barron, J., Karlberg, L., & Rockstrom, J. (2012). Up-scaling potential impacts on water flows from agricultural water interventions: Opportunities and trade-offs in the Osman Sagar catchment, Musi sub-basin, India. *Hydrological Processes, 27*(26), 3905–3921.

Glendenning, C. J., & Vervoort, R. W. (2010). Hydrological impacts of rainwater harvesting (RWH) in a case study catchment: The Arvari River, Rajasthan, India. Part 1: Field-scale impacts. *Agricultural Water Management, 98*, 331–342.

Glendenning, C. J., & Vervoort, R. W. (2011). Hydrological impacts of rainwater harvesting (RWH) in a case study catchment: The Arvari River, Rajasthan, India. Part 2: Catchment-scale impacts. *Agricultural Water Management, 98*, 715–730.

Government of India. (2012). *Final report of minor irrigation and watershed management for the twelfth five year plan (2012–2017)*. New Delhi: Planning Commission, Government of India.

Ho, D. E., Imai, K., King, G., & Stuart, E. A. (2007). Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis, 15*, 199–236.

Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association, 81*(396), 945–960.

Hope, R. A. (2007). Evaluating social impacts of watershed development in India. *World Development, 35*(8), 1436–1449.

Iacus, S. M., King, G., & Porro, G. (2009). cem: Software for coarsened exact matching. *Journal of Statistical Software, 30*(9), 1–27.

Iacus, S. M., King, G., & Porro, G. (2011a). Multivariate matching methods that are monotonic imbalance bounding. *Journal of the American Statistical Association – Theory and Methods, 106*(493), 345–361.

Iacus, S. M., King, G., & Porro, G. (2011b). Causal inference without balance checking: Coarsened exact matching. *Political Analysis, 20*, 1–24.

Joshi, P. K., Jha, A. K., Wani, S. P., Sreedevi, T. K., & Shaheen, F.A. (2008). *Impact of watershed program and conditions for success: A meta-analysis approach. Global theme on agro ecosystems. Report no. 46*. International Crops Research Institute for the Semi-Arid Tropics.

Kerr, J. (2002). Watershed development, environmental services, and poverty alleviation in India. *World Development, 30*(8), 1387–1400.

Kerr, J., Pangare, G., Pangare, V. L., & George, P. T. (2000). An evaluation of dry land watershed development projects in india. *EPTD Discussion Paper No. 68*. Washington, DC: International Food Policy Research Institute.

NABARD. (2006). *Watershed development fund guidelines* (revised as on 31st January 2006) (accessed on 26th July 2013).

Rockstrom, J., Karlberg, L., Wani, S. P., Barron, J., Hatibu, N., Oweis, T., *et al.* (2010). Managing water in rainfed agriculture – The need for a paradigm shift. *Agricultural Water Management, 97*, 543–550.

Sahrawat, K. L., Wani, S. P., Pathak, P., & Rego, T. J. (2010). Managing natural resources of watersheds in the semi-arid tropics for improved soil and water quality: A review. *Agricultural Water Management, 97*, 375–381.

Sharma, B. R., Rao, K. V., Vittal, K. P. R., Ramakrishna, Y. S., & Amarasinghe, U. (2010). Estimating the potential of rainfed agriculture in India: Prospects for water productivity improvements. *Agricultural Water Management, 97*, 23–30.

Shiferaw, B., Reddy, V. R., & Wani, S. P. (2008). Watershed externalities, shifting cropping patterns and groundwater depletion in Indian semi-arid villages: The effect of alternative water pricing policies. *Ecological Economics, 67*, 327–340.

Singh, R., Garg, K. K., Wani, S. P., Tewari, S. P., & Dhyani, S. K. (2014). Impact of water management interventions on hydrology and ecosystem services in Garhkundar-Dabar watershed of Bundelkhand region, Central India. *Journal of Hydrology, 509*, 132–149.

Sreedevi, T. K., Wani, S. P., Sudi, R., Patel, M. S., Jayesh, T., Singh, S.N., et al. (2006). On-Site and off-site impact of watershed development: A case study of Rajasamadhiyala, Gujarat, India. *Global theme on agroecosystems report no. 20*. Patancheru, Andhra Pradesh, India: International Crop Research Institute for the Semi-Arid Tropics.