IFPRI Discussion Paper 02069

December 2021

Evaluating the Impact of Multi-Intervention Development Projects

The case of Ethiopia's Community-based Integrated Natural Resources Management Project

Gashaw Abate

Alan de Brauw

Nick Minot

Rob Vos

James Warner

Solomon B. Wassie

Shijie Yang

Markets, Trade, and Institutions Division

INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE

The International Food Policy Research Institute (IFPRI), a CGIAR Research Center established in 1975, provides research-based policy solutions to sustainably reduce poverty and end hunger and malnutrition. IFPRI's strategic research aims to foster a climate-resilient and sustainable food supply; promote healthy diets and nutrition for all; build inclusive and efficient markets, trade systems, and food industries; transform agricultural and rural economies; and strengthen institutions and governance. Gender is integrated in all the Institute's work. Partnerships, communications, capacity strengthening, and data and knowledge management are essential components to translate IFPRI's research from action to impact. The Institute's regional and country programs play a critical role in responding to demand for food policy research and in delivering holistic support for country-led development. IFPRI collaborates with partners around the world.

AUTHORS

Gashaw Abate (g.abate@cgiar.org) is a Research Fellow in the Markets, Trade, and Institutions Division of the International Food Policy Research Institute (IFPRI), Washington, DC.

Alan de Brauw (a.debrauw@cgiar.org) is a Senior Research Fellow in IFPRI's Markets, Trade, and Institutions Division, Washington, DC.

Nick Minot (n.minot@cgiar.org) is a Senior Research Fellow in IFPRI's Markets, Trade, and Institutions, Washington, DC.

Rob Vos (r.vos@cgiar.org) is the Director of IFPRI's Markets, Trade, and Institutions Division, Washington, DC.

James Warner (j.warner@cgiar.org)is a Research Coordinator in IFPRI's Markets, Trade, and Institutions Division, Addis Ababa, Ethiopia.

Solomon B. Wassie (sbizuayehu@gmail.com) works for Bahirdar University in Ethiopia and is an IFPRI Consultant.

Shijie Yang (syang1@worldbank.org) worked for the Independent Office of Evaluation, International Fund for Agricultural Development (IOE-IFAD) in Italy, at the time of writing. She is currently with The World Bank.

Abstract

This paper provides a quantitative impact assessment of the community-based integrated natural resources management project (CBINReMP) in the Lake Tana region in Ethiopia during 2011-2019. By promoting greater community participation, the CBINReMP provided support to watershed communities for the restoration of degraded soils and water sources, rehabilitation of forests, as well as in obtaining access to secure land titles and practices for climate change adaptation. The project further provided support towards diversification of incomes in off-farm activities and incentives for women's empowerment and youth employment. This way the project aimed to support rural livelihoods through improvements in household incomes, dietary diversity, agricultural productivity, and resilience to climatic shocks, among other livelihood objectives.

To assess the project's impacts, the study had to deal with numerous methodological complications owing to as the project's nature and design. The lack of a proper baseline survey, incomplete information about targeted watershed communities and often lack of clear distinction lines between the project's interventions and support provided to communities through other mechanisms made it hard to identify the true impact of the CBINReMP. Four additional challenges had to be faced: possible selection biases because of non-random placement (targeting) of the project; self-selection of beneficiaries into receiving the project; possible spatial spill-over effects of project benefits to non-treatment communities, and the project's phased rollout. A propensity-score matching procedure was adopted to assess the CBINReMP's impacts by comparing treatment (beneficiary) and control groups outcomes related to the livelihood indicators listed above. This paper discusses how the mentioned complications were addressed to provide a sound assessments of the project's true impacts. While certain limitations remain, the key finding that can be drawn with confidence is that the CBINReMP had only very limited, quantitatively verifiable impact on rural livelihoods. It seems to have contributed to higher household incomes and some greater dietary diversity, but only where the project managed greater community participation. However, even for those beneficiaries, livelihood conditions had not become significantly more productive, diversified, resilient, or sustainable than those of the comparison group. The paper ends with recommendations on how to avoid methodological obstacles through better design of the M&E framework for multi-intervention, community-based projects.

Acknowledgments

This research was conducted for and funded by the Independent Office of Evaluation of the International Fund for Agriculture and Development (IFAD). Co-financing received from the CGIAR Research Program on Policies, Institutions and Markets (PIM) is gratefully acknowledged.

Acronyms and abbreviations

ANRS	Amhara National Regional State
ATT	Average treatment effect on the treated
BoANR	Bureau of Agriculture and Natural Resources
CBINReMP	Community-based Integrated Natural Resources Management Project climate change adaptation
CHIRPS	Climate Hazards Group InfraRed Precipitation with Station Data country strategic opportunities programme
FTC	Farmer training center
GIS	Geographic information system
IFAD	International Fund for Agricultural Development
IFPRI	International Food Policy Research Institute
IOE	Independent Office of Evaluation of IFAD
LASSO	Least Absolute Shrinkage and Selection Operator (regression model)
LTWs	Lake Tana Watersheds
M&E	Monitoring and evaluation
MODIS	Moderate Resolution Imaging Spectroradiometer
NDVI	Normalized Difference Vegetation Index.
NDWI	Normalized Difference Water Index
NIR	Near-infrared
PSM	Propensity Score Method
SLM	Sustainable land management

1. Introduction

This paper assesses the quantitative impact of the community-based integrated natural resources management project (CBINReMP) implemented by the Amhara Region Bureau of Agriculture and Natural Resources (BoANR), IFAD, and partners in the Lake Tana Watersheds (LTWs) in Ethiopia during 2011-2019.[1] By promoting greater community participation, the project provided support to watershed communities for the restoration of degraded soils and water sources, rehabilitation of forests, as well as in obtaining access to secure land titles and practices for climate change adaptation. The project further provided means to communities and households for diversifying incomes through off-farm activities, as well as incentives for women's empowerment and job creation for youth. This way the project aimed to improve livelihoods through improvements in household incomes, dietary diversity, agricultural productivity, and resilience to climatic shocks, among other livelihood objectives.

Several aspects in the design and implementation of the project complicated rigorous, quantitative impact assessment. First, the project involved multiple interventions dependent on community needs and participation complicating the identification of the project's contribution (as a whole and its components) to observed outcomes. Second, the project's targeting of beneficiary communities and households was not clearly defined from the outset, complicating the identification of proper control group to assess "treatment effects". Third, furthermore, not all beneficiary communities received support through the project for the entire range of interventions, while many non-beneficiary communities in the same area received, again in varying degree, support from other agencies for similar interventions, thereby further complicating the identification of control groups and benefits attributable to the CBINReM project. Fourth, many of the beneficiary communities not only already had received support in watershed management prior to the project, but also several project interventions underwent change during the course of its eight-year implementation. For instance, the climate change adaptation component was introduced after a mid-term evaluation. Lacking a baseline and intermediate survey data makes it hard to assess how those changes and their differential implementation weigh on the outcomes. Lastly, the impact assessment also meant to use geo-spatial data to analyse changes in vegetation coverage and other agro-ecological and landscape variables. The use of this type of data was hampered by absence of shapefiles for both beneficiary and control-group watersheds. As a result, the impact assessment for these dimensions could only be undertaken by approximation, as explained further in Section 3.

[1] This impact assessment was part of an overall evaluation of the project undertaken by the Independent Office of Evaluation (IOE) of IFAD, including further a qualitative evaluation of the project's efficiency and effectiveness in achieving its overall goals (IFAD 2021). This part of the assessment refers to an ex-post quantitative analysis of the project's contribution to observed impacts on the livelihoods of beneficiary communities and households. The full evaluation report can be accessed here.

The impact assessment method was chosen to account as much as possible for these attribution problems. Admittedly, not all of these confounders of actual impacts could be fully addressed, as explained in Section 3. Primary survey data were collected among representative samples of beneficiary and control-group households[2] and watershed community leaders after the end of the project (in March 2020).

We find that, overall, the CBINReM project had only very limited quantitatively verifiable impact on rural livelihoods. It seemed to have contributed to higher household incomes and some greater dietary diversity, but only where the project managed greater community participation. However, even for those beneficiaries, livelihood conditions had not become significantly more productive, diversified, resilient, or sustainable than those of the comparison group. Furthermore, it is not clear what factors would have led to greater community participation than in other participant communities, and the contributions noted above might be due to those underlying factors rather than the CBINReM project.

Despite the limited observed impacts and also despite the limitations and attribution problems underlying that assessment, the present analysis should serve policy makers and development partners in Ethiopia and elsewhere regarding how to improve on the design of community-based interventions for sustainable rural livelihoods and poverty reduction, as well as for the design of monitoring and evaluation (M&E) frameworks to facilitate improved impact assessment.

The remainder of this report is organized as follows. Section 2 provides a description of the project's objectives, design and implementation, as well as of its theory of change from which we derive the key questions for the impact assessment. Section 3 describes the community and household survey design, the geo-spatial data used for assessing impacts on vegetation and water retention, and the methodology followed for assessing the impacts on rural livelihoods. This section also explains the matching procedure for comparing treatment and control groups. Section 4 provides a summary of key household and community characteristics of treatment and control groups. Section 5 presents the main findings of the impact assessment, while section 6 concludes with a brief summing up of the key findings, the limitations of this assessments and recommendations for improving project design and monitoring such that it will allow for better impact assessments of future community-based watershed management and rural livelihoods projects.

[2]In the context of this report, the control group refers to the watersheds and households that are located either within or neighbor to the kebeles where there is project intervention at least in one of the watersheds.

2. The Community-based Integrated Natural Resources Management Project

Despite relatively high economic growth over the past decade, the agricultural sector in Ethiopia is still characterized by its subsistence nature and low productivity. Agricultural systems are highly dependent on climate and are vulnerable to extreme climate events. Environmental degradation, as exhibited in significant land and water resource degradation together with biodiversity loss and deforestation, remains a key challenge. Ethiopia loses some 2 billion tons of fertile soils annually to land degradation, and the siltation of water bodies is already a major threat to irrigation development (IFAD, 2009a). Recent estimates using satellite imagery show that land degradation hotspots over the last three decades cover about 23 per cent of the land area in the country (Gebreselassie, Kirui, and Mirzabaev, 2016). About one third of rural households farmed less than 0.5 ha in rain-fed agriculture, which was insufficient to produce enough food to meet the intake requirements of the average household. Most agricultural production was used to meet household consumption needs, and most households experienced a prolonged deficiency in food availability during the pre-harvest period (IFAD, 2009b).

Project area

Against this background, the CBINReMP was designed and implemented by the Amhara Region BoANR, IFAD, and partners in the LTWs in Ethiopia during 2011-2019.The project area comprised the entire Lake Tana Watershed in the Amhara National Regional State (ANRS) in Ethiopia. At the start of the project, about 2.3 million people belonging to 450,000 rural households lived in the area, distributed over 24 woredas (districts) and 347 kebeles (villages). Lake Tana accounts for almost half of total water surface in the country. For decades, the project area suffered from severe land degradation caused by overgrazing, deforestation, unsustainable agricultural practices, and over-exploitation of wetlands. The sustainability of livelihoods in the area further suffered from encroachment on fragile hillsides, insecurity of land tenure, growing population pressure and related increased land fragmentation, and high dependence on bio-mass energy which deprived soil organic matter (IFAD, 2009b).

The project aimed to facilitate land certification benefits to all 450,000 smallholder farm households living in the LTWs area. About 312,000 households were to benefit, in addition, from pasture development, soil and water conservation, participatory forestry, and development of watershed management plans. About 25,000 unemployed youth, including women, would be trained, and assisted in engaging in off-farm income generating activities (IFAD, 2009b). For all beneficiaries, improving incomes was stated as a primary outcome of the project. According to the project design report, per capita income of the target group averaged US$80 per year in 2009, well below the national average per capita income of US$340 and the then national poverty line of US$110 per person per year (IFAD 2009b: vii). Most beneficiary households depended on agriculture for their livelihoods, facing limited access to land and agricultural inputs, severe soil erosion and high vulnerability to impacts of climate change.

Project interventions

The project aimed to address these challenges with a set of interventions promoting sustainable natural resource management with a high level of community participation. Both the role of the community and the nature of the participation of beneficiaries added to the complexity of the impact assessment. First, to understand the project's implementation and outcomes, it is important to note that interventions were planned at the kebele level. Kebeles are the smallest administrative unit in Ethiopia, similar to wards or villages in other contexts. Kebeles typically consist of multiple micro-watersheds. Interventions were targeted towards micro-watersheds but not all watersheds within a kebele were selected as beneficiary communities. This complicates the assessment of the project's impacts, since support coordinated at the kebele level might benefit both targeted and non-targeted watersheds within the kebeles. A second important aspect to note is that community participation included in-kind (labor-time) contributions from beneficiary households to implement actual interventions at the community level, such as for terracing for soil improvement, rehabilitation of watershed gullies, etc., were organized and coordinated at the kebele level. These contributions were made as part of Ethiopia's existing "mass mobilization" program. For assessing the impact, one would not only need to be able to identify whether the project helped leverage additional community labor being provided but also whether such labor inputs were redirected towards the implementation of the project activities.[3] We address this in Section 5.

Community-based integrated watershed management formed the key component of the project's interventions. For this, the project targets were to (a) improve land administration and certification for all rural households in the 21 districts (woredas) of the Lake Tana Watershed; (b) promote participatory watershed planning and management in 13 woredas covering 650 micro-watersheds for a total area of 227,500 ha; (c) establish a database of existing land use patterns and natural resources; (d) improve pasture and forage management in 630 sites covering 9,450 ha of communal grazing lands; (d) rehabilitate 18,900 ha of degraded community forests; (e) establish participatory forest management covering some 10,000 ha in five sites of public forests; (f) rehabilitate 32,500 ha through off-farm soil and water conservation measures; and (g) conserve and enhance biodiversity. In addition to these targets, as mentioned, approximately 25,000 unemployed youth were expected to benefit from off-farm income-generating activities and employment opportunities. Also, at a later stage (from 2016), a component was added which aimed at mainstreaming climate change in the

[3]It is relevant to point out here that, according to the qualitative project impact assessment (IFAD 2021), community leaders did consider that the participation in community works increased the community's sense of responsibility for natural resource use. At the same time, however, in many cases increased participation took the form of providing actual labor time but not necessarily associated with a sense of being able to give direction to planning of the watershed works. Among 24 communities visited for the qualitative assessment, leaders of 23 communities indicated they felt they had little (46%) to only some (50%) influence on the watershed plan, as that planning mainly took place at the kebele level. They described the planning approach as "top down" with government institutions taking decisions that were subsequently communicated to the communities for implementation.

project activities and was articulated into two sub-components, namely: adaptation to climate change and mitigation of climate change. However, this component was implemented only in a limited number of watersheds.

Targeting of beneficiaries

The targeting of beneficiary households was neither systematic nor transparent. According to the project concept report and findings from the qualitative assessment (IFAD, 2021), all residents in the targeted watersheds were considered beneficiaries. Moreover, no information was collected about beneficiaries' characteristics and their baseline performance measured by main outcome indicators, derived from the list of proposed interventions. Instead, the targeting strategy focused on watershed communities assumed to meet the criteria. All households belonging to selected communities were, in principle, considered beneficiaries, though for certain interventions ad hoc criteria were introduced to target beneficiary households within selected communities.[4]

The community targeting also lacked a clear approach. A review of both the project design report and project implementation manual could shed no light on the process of selection of the 650 watershed communities, which were ultimately selected as beneficiaries. The project completion report indicates that the watershed selection was based on the "level of degradation of the watershed, the presence of gullies that are beyond the capacity of smallholding farmers to restore, and *woredas* with no intervention from other projects/donors" (IFAD 2019: 5).However, no complete listing of watersheds existed or could be provided, although – according to the project implementation manual - the ANRS has been said to have identified 800 "micro-catchment areas" belonging to the Lake Tana Watershed (IFAD, 2009c).

The project's theory of change

The project's theory of change consisted of three pathways to achieve its targets. An underlying assumption for the project's successful implementation was that government agencies of the Amhara National Regional State were committed to adopt policy, enforce regulatory frameworks promoting and facilitating landscape-scale watershed management practices (IFAD, 2009b).

The first pathway (and overall outcome) is towards "increases in household income", resulting from all other project outcomes. The second, labeled as "intensification and extensification of river basin management" is premised on ANRS bureaus encouraging and raising awareness among communities of the benefits of improved and participatory resources management in the Lake Tana river basins. Local communities through their watershed management committees were to take greater responsibility for sustainable resource management, seeking both social equity and empowerment of women and youth in the process. The third pathway, labeled "increased resilience of watershed resource users" assumes that with more participatory watershed management improved practices for

[4] For example, the activities facilitating production and use of biogas in some watersheds were targeted at households possessing larger numbers of livestock.

landscape conservation and climate smart agriculture would be widely adopted increasing resilience of users.

Based on the theory of change, the following key questions were raised for the quantitative impact assessment:

- Did the project contribute to better socio-economic conditions among project beneficiaries, in terms of incomes, assets and food security?
    - If incomes increased, was this due to income diversification (e.g. income generating activities) and/or improvement of agricultural productivity?
    - Likewise, did the project improve ownership, or security of access to land, water or productive resources?
    - What helped improve yields? Higher levels of technology adoption, or higher soil fertility, or adoption of suitable agricultural practices?
    - Did project beneficiaries have improved and more regular access to sufficient nutritious food (e.g. as measured in greater dietary diversity)?
- Did the project contribute to strengthen and improve participatory community watershed management practices?
    - Did project beneficiaries increase participation in watershed planning and management?
    - Did the project contribute to empower women's in watershed and household-level decision-making? More in particular, did the land certification empower women in this regard?

Regarding possible "contamination" or "spillover" effects, three important considerations had to be accounted for in this evaluation: (1) the nature of the project intervention where spillover effects may arise; (2) the mechanism through which the spillover effects may occur; and (3) possible confusion of the project intervention with other, complementary types of support organized at the kebele level. In the case of CBINReMP, the land and forest management, watershed restoration, and soil conservation interventions likely created spatial spillovers. These could be a result of ecological processes, but could also be the result from behavioral responses, such as when restricting access to resources in one area induces a rise in extractive activity elsewhere (Baylis et al, 2016; Deininger and Xia, 2016; Ostwald and Henders 2014). Spillovers not only affect net impacts but can also bias impact estimation when they influence non-target areas that were intended to serve as control observations. Both the nature and mechanism of spillover effects influence the impact assessment design and underlying identification strategy, which will be further looked at during methodology development. Regarding the third consideration, based on the qualitative assessment, beneficiaries were not always able to identify the project-specific activities as distinct from other recurrent natural resource conservation interventions. Hence, for the present assessment it had to be assumed that the treatment communities and areas only benefited from the CBINReMP's interventions.

3. Methodology, key hypotheses, and survey design

Methods

The principal aim of this evaluation is to assess the impact of the project on project beneficiaries. Impacts are assessed for four outcomes considered key to rural poverty reduction: (i) increases in household income and assets; (ii) improved human and social capital and empowerment; (iii) improved food security and agricultural productivity; and (iv) strengthened community institutions and participation.

The overall impact evaluation of the CBINReMP conducted by IFAD's Independent Office of Evaluation and IFPRI employed a mixed-method approach. Both quantitative and qualitative data were collected, with the latter being collected prior to quantitative data collection to help inform the design of the quantitative survey. The qualitative data were used to inform interpretation of the quantitative results. Additionally, geo-spatial data were analysed to assess the biophysical indicators as outlined in the theory of change. Here we outline the quantitative approach and the use of geo-spatial data.

This is an ex-post impact evaluation conducted after completion of the project activities. Lacking proper baseline survey data of beneficiary communities and households,[5] a quasi-experimental design method was used to estimate average treatment effects through comparison of beneficiaries and a "control" group.

To evaluate the impact of the project on household income, agricultural productivity, and other social economic indicators, the impact evaluation must attempt to account for potential observable sources of selection bias, with the idea that by accounting for those observables, unobservables are also somehow balanced between the treatment and control groups.

[5] A baseline survey was not undertaken until after several years of the start of the project. The late undertaking of the baseline survey implies that the state of conditions that existed in the project areas prior to CBINReMP interventions cannot exactly be established. Also, as noted in the mid-term review of the project (IFAD 2014), the baseline survey that eventually was conducted in 2013 was not considered to sufficiently comprehensive in design and information coverage to facilitate proper monitoring and evaluation of the project's achievements.

In doing so, the impact assessment had to face the challenges identified in the previous section:

- selection bias because of non-random placement (targeting) of the project;
- self-selection of beneficiaries into receiving the project;
- possible spatial spill-over effects of project benefits to non-treatment communities; and
- a phased rollout approach.

To account for the non-random placement of the project, we control for observable community-level characteristics and geographical attributes that are exogenous to the project – i.e. most of which refer to the period before the project intervention and might be correlated with the project's targeting strategy. However, we acknowledge that the evaluation cannot account for all possible unobservable confounders. In the context of this study, all households living within the targeted watersheds are considered as beneficiaries, so the results can be considered as "intent-to-treat" effects. Hence, self-selection of the beneficiaries to take part in the community watershed activities is not an initial challenge.

As planning of the project intervention was done at the kebele level, the interventions could have benefited both targeted and non-targeted watersheds within a treated kebele. To check for potential spatial "spillover" effect due to the kebele level planning of the project, we first identified whether the control watersheds belonged to a kebele which included a treated watershed or not. We then re-estimated the treatment effects, comparing separately the targeted watersheds with control watersheds located either within or outside the kebeles with treated watersheds. The results of this exercise (reported in Annex Table A.1) do not show consistent pattern that would support the argument of detectable "spillover" effects due to the design of the project.

Lastly, it was not possible to account for any influence of the phased roll out of the project interventions, possessing only after-project information of beneficiary household and community characteristics and overall benefits they received, not how or when they were phased in.

An additional challenge was to identify a proper control group in light of the way beneficiary watersheds were selected. As stated above, the initial selection of watersheds gave priority to those with higher perceived resource degradation. As explained further below, we randomly selected the control group watersheds from a list of non-project watersheds. Since the non-project watersheds thus likely would face less resource degradation this could influence the assessed outcomes, given possible difference in key initial conditions. To account for this potential "mismatch" in conditions between treatment and control group, the household and community survey questionnaires included questions regarding the (perceived) state of natural resource degradation at the start of the project (10 years ago) and this information was used in the matching procedure minimizing such differences.

Survey design

The quantitative data were collected both at the household and community levels. The CBINReMP was implemented in three macro-watersheds covering four zones (i.e. West Gojjam, Central Gondar, South Gondar, and Awi) around the Lake Tana sub-basin. Specifically, the project covered 24 intervention woredas or districts. In two of these woredas, Quarit and Yilmana Densa, only one micro-watershed was targeted and, consequently, had to be dropped from the sample selection. Furthermore, in South Gondar only one component of the project (land certification) was implemented in all five woredas and no information was available for the list of watersheds covered by the project in the kebeles belonging to these woredas. Likewise, three woredas (Wogera, Gondar Ketema, Dangla Ketema) with only either treatment or control kebeles/watersheds were also excluded. Thus, the quantitative impact assessment had to be limited to the 14 woredas for which watershed level information on implementation activities was available. Within these 14 woredas, the project reportedly reached about 153 kebeles and 517 community or micro-watersheds. These kebeles and micro-watersheds constituted the sampling frame for treated or beneficiary watersheds.

A three-stage sampling strategy was followed. In the first stage, three kebeles each from the nine woredas having 10 or more treated kebeles and two kebeles each from the remaining five woredas, with less than 10 treated kebeles, were selected using simple random sampling. Thus, a total of 37 treated kebeles were considered. In the second stage, two treatment watersheds were selected from each sample kebele selected in the first stage using simple random sampling. The sample of watersheds was drawn from the list of watersheds initially targeted by the project. In the third stage, based on the list of community members provided by the watershed management committee, 12 farm households were selected from each community watershed using systematic random sampling.

Once the sample treatment kebeles were identified, it was decided to select control group community watersheds and households from a list of non-intervention kebeles neighbouring to the selected treatment kebeles. This decision was made on grounds of similarities in agro-ecological conditions and presumably also socio-economic conditions. While this could not be fully verified during the sampling process, it was further assumed that the control group kebeles and watershed communities not only had no part in the CBINReMP but also not from any other watershed development project by development partners (other than the periodic natural resource conservation implemented by the government through mass mobilization).[6] The attempt here was to avoid any problem of contamination of intervention benefits between treatment and control group, while having a proper control group would allow for proper estimation of treatment effects. Following the establishment of the sample frame for control group communities, the same three-stage sample selection procedure was followed for the control group sample selection.

---

[6] Kebeles and watersheds receiving benefits from interventions by other projects with similar objectives to those of the CBNReMP were excluded from the sampling frame, regardless of their treatment status.

The sample size thus obtained consisted of 74 treatment watershed communities and 887 treatment households and 62 control group watershed communities and 768 control households (Tables 1 and 2).

Table 1. Survey sample design and distribution between treatment and control groups

Description	*Treatment group*	*Control group*	*Total sample*
Number of woredas	14	14	28
Number of kebeles	37	31	68
Number of watersheds	74	64	138
Number of households	887	768	1,665

Table 2. Geographic distribution of the sample by treatment and control group

Zone	Woreda	Number of watersheds		Number of households	
		Control	Treated	Control	Treated
West Gojjam	Bahirdar Zuria	6	6	72	72
West Gojjam	Bahirdar Ketema	4	4	48	48
West Gojjam	North Mecha	4	4	48	47
West Gojjam	South Mecha	4	4	48	48
West Gojjam	Sekela	4	6	48	72
West Gojjam	North Achefer	6	6	72	72
West Gojjam	South Achefer	4	6	48	72
Awi	Fagitalekoma	6	6	72	72
Awi	Dangla Zuria	6	6	72	72
Awi	Banja	4	4	50	48
Central Gondar	Gondar Zuria	4	6	48	72
Central Gondar	West Dembia	4	4	48	48
Central Gondar	East Dembia	4	6	48	72
Central Gondar	Lay Armacheho	2	6	46	72

Questionnaires and survey implementation

Household and community questionnaires were developed, pre-tested in the field, and modified accordingly before the actual survey data collection, which took place during March 2020. Of the1,674 households identified from the sampling frame for interview, 1,665 of them were available and willing to complete the household survey implying a response rate of 98.9 percent. Likewise, community level data was collected from 136 sample micro-watersheds. One key informant (typically head of household) was interviewed for collecting the household-level data, while several respondents were sought to provide the information relating to the community survey questionnaire (typically, two members of the community watershed committee, one or two elders from the community, and woman and youth representatives).

The questionnaire of the community survey included questions regarding community organization; community's access to infrastructure, institutions, services, and markets; community-led natural resource conservation and climate adaptation practices. The household survey included modules on household composition, land use, land certification, crop and livestock production and utilization, natural resource conservation, extension services and credits, off-farm income, food security, adaptation strategies, and participation in community planning and works. Interviews were conducted in Amharic, local language of the study area.

Geo-spatial data

This impact assessment makes use of agro-climatic and geo-spatial data to assess the biophysical indicators as outlined in the theory of change. According to the project design report, interventions for all targeted 650 watersheds were designed using geo-spatial information. However, none of the area shapefiles needed to geographically identify micro-watersheds could be provided by the project managers or local authorities.

Due to the unavailability of the shapefiles, new watershed area data were created. The total sampled watershed area was 're-created' from information provided by respondents to the community questionnaire; specifically, using the responses to the questions regarding how much time it took, in minutes, to walk from the north to the south edge, as well as from the east and west edge. This walking time was converted to distance and then projected into an estimated rectangle area of the watershed. The GIS-derived centroid was then applied to centre of the rectangle. On this basis, we estimated that the mean of the sampled watershed area was 7.7 km^2 with a median of 5.2 km^2. Given the application of a uniform walking time, imposed boundary form and typical variations in respondent estimation, these estimates should be taken with a fair degree of possible error. For instance, although watersheds should be discrete objects, many watersheds had overlapping boundaries or centroids that did not seem to conform to topography. This has implications for treatment and control groups since they were subsequently modelled, in some instances, as overlapping. Regardless of these limitations, remote-sensed data was derived from these rectangles and consists of four major variables.

To capture changes in the landscapes due to interventions, we utilize satellite remote-sensing images from MODIS, LandSat, and a derived dataset called Climate Hazards Group InfraRed Precipitation with Station data (CHIRPS). Spatial datasets are derived from three primary sources both of which were available near the year of the start of project interventions. We use MOD13Q1 and MYD13Q1 MODIS products to construct an interpolated 8-day equivalent Normalized Difference Vegetation

Index (NDVI) time series with a 250m resolution.[7] Landsat 8 collection tier 1 was used to generate annual cloud-free median NDVI. The NDVI is generated from the Near-infrared (NIR) and Red bands of each scene as (NIR - Red) / (NIR + Red), and ranges in value from -1.0 to 1.0. NDVI is sensitive to the presence of chlorophyll and is regularly used as a proxy for plant health and productivity. From the same source, we also calculate annual Normalized Difference Water Index (NDWI) which is sensitive to changes in water content of vegetation, with values ranging from -1 to 1.[8] Both LandSat products are annual but have a significantly higher spatial resolution than MODIS products (30m versus 250m, respectively) The time series properties of rainfall are measured by the CHIRPS dataset. CHIRPS incorporate 0.05° resolution satellite imagery with in-situ station data to create gridded rainfall time series for trend analysis and seasonal drought monitoring. In this case, we resample the rain data to 75m spatial resolution to ensure that each enumeration area has an observation associated with it. Precipitation is collected by dekad (Funk et al. 2014). There are three dekads in a month, the first two being 10 days long, and the third being the remaining days in the month.

All data is summarized over time to help differentiate changes within treatment and control watersheds. For instance, we might look at whether NDVI or "greenness" is higher in intervention areas than in the control group. The challenge then is to create a set of indicators that meaningfully describes differences between the watersheds for the seven years for which we have data.

A large number of potentially important time-series features were derived from the remotely sensed imagery. For the sake of brevity, we only describe those features that were used in the final analysis. Note, that most time-series indicators will be more robust for the MODIS and CHIRPS because of their significantly higher temporal frequency. In Table 3, we describe the set of metrics extracted and a brief description of each. Each time-series metric described below is then summarized by their mean value for all land within each of the treatment and control watersheds.

Table 3. Description of remote sensed variables (2013-2019)

Name	Description	Interpretation
NDVI/NDWI Slope	Univariate time-series regression estimate	Time trend (positive increasing—negative decreasing)
NDVI/NDWI Standard Deviation	Distribution of observations from mean	Are variations of cropping patterns (water retention) larger/smaller?
Mean	Global mean value	Average observed greenness / rainfall (annual)
Median	Global median value	Median of observed greenness
PPT sum (annual)	Total annual rainfall during the meher crop season	Measures relative rainfall variation

[7] https://lpdaac.usgs.gov/products/mod13q1v006/

[8] https://www.sciencedirect.com/science/article/abs/pii/S0034425796000673

As the data captures the entire watershed and does not allow for spatial heterogeneity within the watershed (i.e., individual plots), our statistical analysis is restricted to statistical differences contrasting treatment and control watersheds. Owing to these limitations, the geo-spatial data are used to provide complementary, contextual information to interpret the results of the quantitative impact assessment based on the household survey data but could not be directly used for the estimation of the treatment effects.

Matching procedure

The propensity-score matching procedure controlled first for initial heterogeneity between watersheds and households, based on the probability of a watershed and household participating in CBINReMP conditional on its observable co-variates. Subsequently, to estimate the treatment effects, a *doubly robust estimation method* was applied, which combines propensity-score estimation and regression-based methods (PSM Weighted Regression) (Wooldridge, 2007). The doubly robust estimation method allowed the evaluation to better account for the observable community characteristics that are correlated with program participation and the outcomes, while assuming that unobservables are also balanced between the participants and control group on average.

Matching was conducted at two levels. The first step consisted of matching treatment and control groups at the watershed/community level**.** Since each *kebele* was assumed to include a pool of qualified micro-watersheds and households possessing similar characteristics as those of project communities and households, the community-level propensity score was adopted to find counterfactual communities outside the project area but either within the same *kebele* or a control watershed from neighbouring *kebele*. A restriction was applied to the communities within the same district to assure geographical similarity and spatial proximity between project watersheds and potential control watersheds. Matching parameters were derived from the community-level data.

Selection of the matching variables was done with due caution, because if the project's objectives were met, some of the variables might have changed because of the project. For example, even the household demographics may also not be valid matching parameters, like marriage or migration. Since CBINReMP was a nine-year project, the project might have affected virtually any variable one could think of at the household level, including variables that are often used in matching models such as household demographic characteristics, asset holdings, or production variables. Therefore, it was decided instead to use variables measured in the community survey that largely reflected pre-treatment variables that could be measured. Since the community or watershed level was the targeted unit of intervention, it made most sense to also develop propensity scores at that level. Ideally, those variables should reflect the type of characteristics used for the selection of beneficiary watersheds for the CBINReMP in the first place. After controlling for these variables, the remaining variation in characteristics of watersheds should be considered to be approximately random, rather than due to unobservable differences between selected and control watersheds.

The variables for the matching of treatment and control group cases were subsequently selected using the LASSO regression model (e.g., Tibshirani, 1996). The LASSO model is a method for selecting variables to be included in a regression in a way that it maximizes predictive value. Intuitively, it is not very different from a standard regression, but with the main difference being that it includes a penalty function for inclusion of variables that do not help explain the outcome. For measuring propensity scores, we combine the LASSO regression with a logit model, in which we use a cross-validation algorithm to choose variables to include in propensity-score estimation. The list of potential variables included community variables that were arguably exogenous, as well as interactions between variables that were continuous or discrete and continuous. The LASSO is increasingly used in studies requiring estimation of propensity scores, particularly in epidemiology. In that literature, Franklin et al. (2015) find that the LASSO outperforms other estimators.

The second step was to use the propensity scores to estimate the predicted probability of inclusion for each watershed. For each household in a watershed, the propensity score indicates the predicted probability that the household belongs to a treated watershed community rather than to a comparison group of non-treated watersheds. The propensity scores p are then used as weights for the comparison observations, that is, while each treatment observation receives a weight of one, the control-group observations receive a weight of $\frac{P(X)}{1-P(X)}$. The intuition is as follows. Watersheds that have observable characteristics indicating that they are not likely to be chosen as participants receive very low weights in the regression, whereas observations with observable characteristics suggesting that they should be good comparisons to treatment observations receive a great deal of weight. By placing higher weights on non-recipient observations that have characteristics more like participants and lower weights on non-participants that have characteristics less like participants, we balance observable characteristics between participants and non-participants, even if they were unbalanced before weighting. Using the weights, next a balance test among observable characteristics—both those included in the propensity score estimation and those that were not—will be conducted to ensure that observable characteristics are balanced after applying the weights based on propensity scores. Details on the variables included in propensity scores and a balance table for observables prior to treatment are included in Appendix Table A.2.

Testing for treatment or degree of participation

The project implemented wide range of activities focusing on participatory watershed management, pasture and forage development, soil and water conservation, and biodiversity and ecosystem protection. However, evidence from the qualitative assessment shows that the degree of participation in the various project activities varied considerably across targeted watershed communities. A descriptive analysis of the participation variables of the household and community surveys also confirmed this was clearly the case. This leads us to make a distinction between "high" and "low" community project participation and assess potential impact heterogeneities. The distinction was made based on close examination of responses to 18 survey questions related to household and community participation in the planned activities of the project (Appendix Table A.3). A "participation score" (ranging from 0 to 18) was created to rank communities from low to high level of participation. To ensure a comparable counterfactual, two of the control-group watershed communities with a participation score of more than 12 were dropped from the sample. The high project participation score in these cases could reflect that, despite being identified as non-treatment, these were nonetheless direct or indirect beneficiaries and hence cannot be considered part of the control group.

In the analysis of the treatment effects (section 5) we make this distinction between "high-" and "low-participation" treatment groups based on the degree of project-related activity participation. Since community participation was both a means to the outcomes and an (intermediate) objective of the project, the distinction made could confound the actual impacts of the project. Based further on the information provided by communities during the qualitative focus group discussions, we interpret higher participation as synonymous to the intensity of the project's effort (i.e. participation level in the treatment) and that more treatment would more likely help generate the targeted outcomes. We return to this issue in Section 5.

Estimation procedure

The quantitative impact assessment will be based on an estimation of the average treatment effect on the treated (ATT) for the projects targeted outcomes. The ATT is estimated as the difference between the outcome variable for the households among which the treatment was administered, and among households that were not offered the treatment. The average treatment effect of CBINReM project is estimated using a doubly robust method, as indicated above when discussing the LASSO method for the matching procedure. That is, while we regress the outcome variable over the treatment status, higher weights were given to non-beneficiary observations with characteristics more like beneficiaries and lower weights otherwise.

Formally, the specification of the regression model used to estimate the ATT can be formulated as follows:

$$Y_{ji} = \alpha_{1j} + \beta_1 \mathrm{Treat}_{ij} + \beta_k Z_{ij} + \varepsilon_{ij} \tag{1}$$

Where: Y is (lead and intermediate) output variable; $Treat$ refers to the treatment status (i.e. –treated or control) which is a measure of treatment effect; Z_1refers to different community-level co-variates selected by the LASSO model; α_1, $\beta_1$and β_k are parameters to be estimated; subscript i denotes households, j indexes watersheds, and k denotes the co-variates; ε is a mean–zero error term. Here, the primary null hypothesis to be tested is whether β_1 (ATT) is equal to zero.

Section 5 presents the main findings from the application of the estimation procedure. The findings are shown not only for the comparison between entire treatment group and the control group, but also for respectively, the high-participation and low-participation treatment groups, on the one hand, and the control group, on the other.

4. Descriptive data

Household and community socio-economic characteristics

By the nature of the project, treated and control groups were not allocated randomly. Hence, to evaluate the extent to which the two groups will be comparable, we executed a series of balancing tests on household and community level characteristics. Accordingly, Table 4 describes the household characteristics of the treated and control groups. The results show that, with the exception of distance to cooperatives, the two groups show neither detectable nor statistically significant differences in their demographic characteristics, asset holdings, and access to training and market centers.

Table 4. Household-level characteristics by treatment status

Variable	Definition and measurement of the variable	Treated group	Control group	Adjusted Wald test
Age	Age of the household head	49.08	49.08	0.01
Education	Education level of the household head	1.65	1.54	0.41
HH size	The number of active labor force in the family	5.72	5.69	0.06
Land holding	Total land owned (ha)	1.29	1.25	0.21
Livestock	Total livestock of the household measured in Tropical livestock unit (TLU)	5.33	5.40	0.09
Distance to FTC	Distance from home to the farmer training center (FTC) using usual means of transport (travel time in minutes)	32.35	35.13	0.78
Distance to woreda center	Distance from home to the Woreda center using usual means of transport (travel time in minutes)	108.4	112.7	0.28
Distance to the cooperative	Distance from home to the cooperative using usual means of transport (travel time in minutes)	42.57	52.97	4.10**

Source: Own computation, 2020

Note: ** refers to 5-percent significance level.

The balancing tests on community level characteristics of the treated and control watersheds are presented in Table 5. The two groups face similar agro-ecology conditions and degrees of access to basic infrastructure and services, such as telecommunication, electricity, and health services. The two groups are also comparable in their total population and area coverage. While, on average, the treated watersheds are located closer to both markets and cooperatives, the treatment and control group communities do not show detectable differences in access to roads and training centers. Overall, though, we conclude that the two groups are comparable for all community level co-variates presented in Table 5.

Table 5. Community-level characteristics by treatment status

Variable	Definition and measurement of the variable	Treated group	Control group	Adjusted Wald test
Distance from Woreda	Distance of the watershed from woreda center (km)	18.75	19.19	0.01
Road access	Distance from the nearest gravel road (km)	2.68	3.04	0.39
	Distance from the nearest asphalt road (km)	17.82	18.06	0.01
Distance to market	Distance from the center of the watershed to the nearest market (km)	5.58	8.99	4.90**
Distance to cooperatives	Distance from the center of the watershed to the nearest cooperatives (km)	4.48	7.30	3.39*
Distance to FTC	Distance from the center of the watershed to the nearest farmer training center (FTC)(km)	2.38	2.69	0.47
Agro ecology	Percentage of lowland agro ecology	4.42	9.59	1.75
	Percentage of midland agro ecology	86.40	79.33	1.37
	Percentage of highland agro ecology	9.17	11.07	0.14
Access to telecom	= 1 if there is access to telecommunications (% with access)	86.47	83.62	0.21
Access to electricity	= 1 if there is access to electricity (% with access)	21.64	11.59	2.51
Access to health center	= 1 if there is access to health center (% with access)	44.64	37.79	0.64
Population	Total number of households in the watershed	256.0	300.0	2.43
Area	Total area of the community watershed (ha)	433.3	452.4	0.24

Source: Own computation, 2020

Note: * and ** refer to 10 and 5percent significance level, respectively.

Geo-spatial characteristics

Four spatially derived variables were used to assess whether control or treated watersheds exhibited important differences regarding vegetation cover changes, soil water retention mapping (irrigation or other water management strategies) or were impacted by relative annual rainfall differences. Given that the data was not normally distributed, median tests were performed. Table 6 indicates that none of the variables were found to be statistically different, suggesting geo-spatial conditions were roughly similar on average for the watershed areas where the control and treatment groups were located. However, it should be remembered that the lack of clearly delineated, mutually exclusive, boundaries implies this conclusion has to be taken with great caution.

Given this caveat, NDVI and NDWI trend lines were drawn through the data to determine if there were changes in vegetation coverage over the 7-year period of observation (2013-2019). A positive slope would imply increase greening of the watershed over time, while a negative slope would indicate a deterioration of vegetation cover. While both the MODIS and Landsat harvested variables revealed a statistically significant positive slope for the median of the sampled watersheds, there were no statistical differences between the treatment and control groups. The potential reasons for the overall positive slope could be attributed to improved erosion techniques or common land rehabilitation undertaken in all watersheds but it may also be due to exogenous factors like increased rainfall experienced during the final years of the project's implementation. The median water index was slightly negative with no statistical differences between the two groups (it should be noted that the overall mean was slightly positive because of a few large positive values).

Table 6. Geo-spatial characteristics by treatment status

Variable	**Definition of the variable—Time (2013-2019)**	**Control group (median)**	**Treated group (median)**	**Wilcoxon rank-sum test (Mann-Whitney)[a]**
NDVI_MODIS_slope	Univariate regression slope of Modis NDVI	.0004	.0004	0.88
NDVI_LS__slope	Univariate regression slope of Landsat NDVI	.0027	.0024	0.77
NDWI_LS__slope	Univariate regression slope of Landsat NDWI	-.0013	-.0013	0.97
NDVI_MODIS__sd	Modis NDVI (standard deviation)	.1528	.1521	0.94
NDVI_LS__sd	Landsat NDVI (standard deviation)	.0541	.0534	0.60
NDWI_LS_ _sd	Landsat NDWI (standard deviation)	.0379	.0384	0.70
NDVI_MODIS_mean	Global Mean NDVI Value	.5388	.5416	0.65
NDVI_MODIS_median	Global Median NDVI Value	.5385	.5407	0.66
PPT_sum_2013	Precipitation during 2013 meher crop season (cm)	1,365	1,424	0.66
PPT_sum_2014	Precipitation during 2014 meher crop season (cm)	1,335	1,317	0.68
PPT_sum_2015	Precipitation during 2015 meher crop season (cm)	1,260	1,260	0.78
PPT_sum_2016	Precipitation during 2016 meher crop season (cm)	1,252	1,248	0.77
PPT_sum_2017	Precipitation during 2017 meher crop season (cm)	1,518	1,500	0.67
PPT_sum_2018	Precipitation during 2018 meher crop season (cm)	1,324	1,305	0.67
PPT_sum_2019	Precipitation during 2019 meher crop season (cm)	1,391	1,390	0.60

Source: Own computation, 2020

Note: No statistically significant differences were found. [a] The Wilcoxon rank sum test is a non-parametric test that may be used to assess whether two distributions of observations obtained between two separate groups on a dependent variable are systematically different from one another.

We subsequently looked at changes in geo-spatial conditions over the 2013-2019 period by testing standard deviations for the key indicators. Again, we did not find statistically significant differences between control and treatment groups. Given that the MODIS product was collected at a higher frequency (every 8 days versus an annual aggregation for Landsat), further tests on the means and medians were performed, but also in this case no statistically significant differences could be identified. Annual variations in rainfall could suggest important variations in NDVI and NDWI indexes but, while there were some annual differences in area rainfall, co-variation suggests relatively similar impacts on both treated and control watersheds.

As indicated, the analysis described above suffers from several limitations. In the final section, we provide some recommendations for a methodology that would overcome such limitations in the use of geo-spatial data for future impact evaluations of this kind.

5. Impact assessment results

Findings of the impact of previous soil and water conservation efforts on land productivity in Ethiopia show mixed results. A study in Tigray region through a 500 households survey suggests that plots with stone terraces experience higher crop yields (Pender and Gebremedhin, 2006). Similarly, Holden et al. (2009) used nearest neighbour and kernel matching to measure the impact of stone terraces in Tigray region and found a significant and positive effect on land productivity. Kassie et al. (2007) had found similar positive findings in another evaluation. Kassie et al. (2009), in contrast, found that plots with bunds resulted in lower yields compared to non-conserved plots using matching methods and switching regression analysis on farm-level data from high rainfall areas in western Amhara. A recent impact evaluation from a watershed intervention program in Ethiopia found that sustainable land management (SLM) practices contributed to water security for both crop and livestock production, as reflected in crop yields for maize, mango, and millets using data for 561 households and 2,900 plots belonging to four watersheds (Kato et al. 2019).

Existing evidence on the impact of watershed developments on livelihood and income improvements are also mixed and based on case studies of limited number of watersheds. For instance, Gebregziabher et al. (2015) found that watershed management has improved farm income and food security considerably but based on a case study of six watersheds. Likewise, Siraw et al. (2017) indicated that community watershed development interventions improved livelihood assets, based on evidence from three watersheds in the northwestern highlands of Ethiopia. At the plot level, a study by Schmidt and Tadesse (2012) indicated a positive impact on livelihoods following the use of sustainable land and water management practices (i.e. terraces, bunds, check dams), but this impact was found only for plots with conservation infrastructure maintained over a period of at least seven years.

Bearing these mixed experiences in mind, the remainder of this section presents estimated average treatment impact of the CBINReM project on lead and intermediate outcome variables. Definitions and measurement units of the outcome variables described in the Tables 7, 8 and 9, presented below, can be found in Appendix Table A.4, while more detailed information on the means and skewness in key outcome and intermediate variables can also be found in the Appendix (Tables A.5a-b and A.6a-b).

Impact of the CBINReM project on main outcome variables

Table 7 shows the impact of CBINReM project on rural livelihoods in terms of the main targeted outcomes of improved household incomes, food security, asset holdings, agricultural productivity, and social capital. The first column of each table compares the average treatment effect between treated and control groups, assuming that there is no significant difference in extent of participation among beneficiaries within the treated watersheds. After relaxing this assumption, the second and third columns show the estimated treatment effects after comparing, respectively, the high- and low-participation treatment groups with the control group.

Overall, the results in Table 7 show no detectable differences between treatment (when taken as a whole) and control groups with respect to livelihoods status, social capital, and agricultural productivity. However, we do observe some significant treatment effects when comparing the "high-participation" treatment group with the control group (column 2 in Table 7). Beneficiary households with high community participation have significantly higher income and greater dietary diversity than the control group. Specifically, the incomes of high-participation treatment households were, on average, 17.8 percent higher than that of the control group.

The dietary diversity score exceeded that of the control group by 0.4 units. Dietary diversity is especially important among populations with diets based on starchy staples where micronutrient deficiency is more likely, as is the case in the project area. A higher score is an indicator of increased economic access to a varied diet for household members (though the indicator does not reflect intra-household dietary patterns) (Swindale and Bilinksy, 2006). While this does not follow directly from the method applied here, likely the better access to more diversified food is closely associated with the higher incomes of the high-participation treatment group.

Finally, we further find a significant and positive treatment effect for cow milk productivity among high-participation beneficiaries. However, we do not find discernable differences for cereal yields or any of the other livestock productivity outcomes.

Table 7. Average impact of the project on lead outcome variables

Outcome variables	Treated vs Control (std. err.)	High-participation Treated vs Control (std. err.)	Low-participation Treated vs Control (std. err.)
A. Livelihood outcomes			
Total income (ln)	0.044 (0.09)	0.178 (0.11)*	- 0.152 (0.12)
Dietary diversity	0.197 (0.16)	0.414 (0.167)**	- 0.110 (0.18)
Food security	- 0.086 (0.24)	- 0.155 (0.26)	0.013 (0.28)
Asset holding	- 0.035 (0.16)	0.062 (00.17)	- 0.173 (0.17)
B. Social cohesion and capital			
Social cohesion index	0.068 (0.17)	0.032 (0.18)	0.128 (0.21)
C. *Agricultural productivity*			
Cereal yields			
White teff yield (ln)	-0.075 (0.09)	- 0.039 (0.09)	- 0.131 (0.11)
Black teff yield (ln)	0.067 (0.09)	0.125 (0.11)	- 0.038 (0.12)
Maize yield (ln)	- 0.069 (0.10)	- 0.089 (0.12)	- 0.052 (0.12)
Livestock productivity			
Lactation period (ln)	0.015 (0.04)	0.034 (0.04)	- 0.013 (0.04)
Milk cow productivity (ln)	0.042 (0.20)	0.084 (0.03)**	- 0.021 (0.04)
Fattening period (ln)	- 0.070 (0.12)	- 0.065 (0.12)	- 0.068 (0.13)

Source: Own computation, 2020

Note: ** and * refer to5 and10percent significance level, respectively. [a] The social cohesion index is a composite of five perceptions about belongingness of individuals in the community regarding economic opportunity, opportunity in public affairs, tolerance to conflict of interest, and adequate representation in institutions. The related questions are contained in module H (questions H2) of the community questionnaire (see Annex 1). In the questionnaire the response categories are: 1 Strongly Agree, 2 Agree, 3 Disagree, and 4 Strongly Disagree. In the estimation of a reverse valuation was used, such that a higher index means higher social cohesion.

In summary, project beneficiaries in communities with high degrees of participation in community-based natural resource management activities enjoy higher incomes and this may also allow them to have better diets. However, these positive livelihood outcomes have not come with other targeted livelihood improvements (relative to the comparison group) in terms of agricultural productivity, social cohesion, or asset holdings. The higher milking cow productivity likely underpins a modest part of the estimated income impact and, while notable, the impact was not among the central targeted outcomes of the CBINReM project. Next, we turn to the assessment of the project's impacts on intermediate outcome variables.

Impact of the CBINReM project on intermediate outcome variables

a. Community participation

A first important intermediate outcome of the project is that it seems to have significantly increased participation of beneficiary households in providing labor time for most of the community works promoted (Table 8). The survey results show that the treatment groups spend visibly more time on communal terrace construction, cut-off drainage and tree planting, though this is not the case for gully rehabilitation. The labor participation in these types of communal works among the 'high-participation' treatment group households is broadly the same as that for the average treatment group. However, the confidence level for all of these estimates is low, such that none of the differences between control and treatment groups were found to be statistically significant. A significant impact for labor participation would have been important in terms of the project's theory of change, which saw enhanced community participation for sustainable land and water management as key to create better and more resilient livelihoods for the beneficiary population. Given the lack of statistical significance we cannot confirm with any certainty that the project was effective or not in promoting community participation in SLM works to underpin livelihood improvements.

For other participatory variables we find little difference between treatment and control groups. For instance, 68% of households of both control and treatment groups participate in the watershed planning process and almost equal shares form part of grazing groups and other forms of community participation (see Appendix Table A.7). Beneficiary communities are somewhat more likely to have a watershed plan (86%) compared with the control group (77%). As indicated in Section 3, however, we did find significant differences in degrees of participation, such that we separated the treatment group in terms of high- and low-participation treatment and participation. Thus, when assessing the project's impact, we can no longer see differences in participation as a true impact for these sub-groups of beneficiaries for obvious endogeneity problems. Hence, the finding shown in Table 9 that the substantially higher shares of households in the high-participation treatment group participating in watershed planning (95% versus 77%) and forming members of shared grazing land communities (61% versus 46%) cannot be taken as intermediate outcomes of the project, but just the confirmation of the relevance of the distinction made.

Table 8. Labor time spent on project-related community works (hours per year)

Type of community work	Control group (A)	Treatment group (B)	High-participation treatment group (C)	Adjusted Wald Test B-A	Adjusted Wald Test C-A
Terrace construction	85	103	108	0.87	1.27
Cut off drainage	37	60	62	2.01	2.23
Gully rehabilitation	42	38	39	0.07	0.04
Tree planting	33	276	293	1.09	1.08

Source: Table A.6a-b.

Table 9. Other key participation variables

	Control group	Treatment group	High-participation treatment group
Participate in watershed planning	77%	86%	95%
Membership in grazing land	46%	51%	61%

Source: Table A.7.

Overall, we cannot take these findings regarding participation as confirmation or rejection of the effectiveness of the theory of change, since we cannot establish causal relationship between this element of community participation and the impact on livelihood improvements discussed above.

b. Project's impact on intermediate variables

We also find few discernable differences between treatment and control group household for a range of key intermediate outcome variables as targeted by the project. For the group of treatment households as a whole, we only find a significant impact on perceived increases in herd size (compared with ten years ago) and this perceived impact is even stronger for the high-participation treatment group (Table 10). While there are no other significant impacts for the treatment group as a whole, we find three other significant impacts for the high-participation treatment group. Specifically, these are benefits from "cut and carry" of forage from communal grazing lands, and (perceived) increased participation on soil and water conservation practices on both own and communal lands.

Regarding the first benefit, since villagers sell at least part of forage from the cut-and-carry practice in the market, this may have contributed to the positive income effects for the high-participation treatment households and likewise for the perceived increase in herd size. Again, however, the estimates in Table 10 do not allow inference of such a causal relationship. Most likely though, the impact on raising average household income has been limited. Although the "cut-and-carry" system has a potential of ensuring sustainable forage off-take, the cutting is done only once or twice per year. Such long cutting return periods may satisfy those who have land or other means of livelihoods, but not the landless or the marginal farmers who have no other alternative means of livelihoods while they are waiting for their biennial cut-and-carry share. Those who have land could use crop residues in the meantime or they may have other ways of producing fodder at the farm.

Regarding the benefits from increased soil and water conservation practices, the qualitative survey (IFAD, 2021) concluded for a sub-sample of beneficiary watershed communities that the project "successfully promoted the construction of physical and biological soil and water conservation (SWC) structures in off-farm degraded areas and supported improved land administration and certification." The finding below of perceived improvements in engagement of SWC practices would be consistent with that finding. However, that assessment recognized, at the same time, that focusing only on off-farm structures did not sufficiently address the causes of land degradation while the project's support to the generation of new income-generating activities little effective in diversifying livelihood opportunities. These implementation weaknesses could explain the lack of impact on key lead and intermediate outcomes, such as crop productivity, income diversification, and resilience to climate change.

Table 10 Average impact of the project on intermediate outcome variables

Outcome variables	Treated vs Control (std. err)	High-participation Treated vs Control (std. err)	Low-participation Treated vs Control (std. err)
A. Resource management: animal feed, conflict and climate change			
Cut and carry	0.022 (0.03)	0.131 (0.03)***	- 0.133 (0.04)***
Free grazing (ln)	- 0.176 (0.27)	-0.207 (0.28)	-0.119 (0.30)
Resource conflict	0.014 (0.026)	0.028 (0.03)	- 0.005 (0.03)
Perceived improved resilience to climate change (wrt 10 yrs ago)	0.066 (0.07)	0.094 (0.08)	0.026 (0.08)
B. Income diversification and women's empowerment			
Income diversification	- 0.009 (0.05)	0.075 (0.06)	- 0.129 (0.06)
Off farm income	0.124 (0.08)	0.128 (0.08)	0.118 (0.08)
Women's participation in watershed committee	1.787 (2.10)	1.882 (2.19)	1.626 (3.26)
C. Perceived improvement in crop and livestock productivity [a]			
Agricultural productivity (10 yr)	0.053 (0.09)	0.126 (0.10)	- 0.051 (0.10)
Cereal yield (10 yr)	- 0.04 (0.10)	- 0.071 (0.11)	0.006 (0.11)
Herd size (10 yr)	0.072 (0.03)**	0.122 (0.03)***	0.002 (0.03)
D. Perceived increase in soil and water conservation participation [a]			
SWC on own land (10 yr)	0.026 (0.04)	0.136 (0.04)***	- 0.118 (0.04)***
SWC on common land (10 yr)	0.033 (0.03)	0.126 (0.04)***	- 0.087 (0.04)**

Source: Own computation, 2020

Notes: *** and ** refer to, respectively, 1 and 5 percent significance level. [a] Refers to share of households reporting improvements compared with 10 years ago.

The qualitative survey (IFAD, 2021) further found that the project helped rehabilitate and protect the vegetation coverage of degraded land in a sub-sample of beneficiary watersheds through the SWC practices. However, as already pointed out in section 4, our geo-spatial analysis could not find significant differences in vegetation coverage or soil water retention between the beneficiary and control group communities. Hence, while the project may have contributed to NDVI improvements, we cannot conclude it was more effective than other interventions and factors that may have helped similar improvements in non-beneficiary watersheds.

For other elements of the theory of change we do not find any significant impacts of the project, such as for income diversification, off-farm income generation, women's empowerment (e.g. through participation in watershed communities), reduction in free grazing land, reducing conflict over land, or – as mentioned – perceived improvements in resilience to climate change or cereal yields.

6. Conclusions

Summary of overall findings

This study set out to assess whether the CBINReMP led to increases in household income (including through income diversification); improved access to land, water and productive resources; improved crop and livestock yields; and greater food security and dietary diversity. These improvements would result from the implementation of sustainable landscape, soil and water conservation practices through community participation. The related interventions were further expected to result in greater resilience among communities and households, including to climate change.

Lacking a proper baseline study, no rigorous assessment could be made of the improvements over time and how much the project's interventions contributed to such interventions. Due to this limitation, the analysis was restricted to an ex-post impact assessment comparing beneficiary households and communities with a comparison group. Doing so, was constrained by further complications and limitations as noted in Section 3. Nonetheless, a few key conclusions can be drawn:

(a) The project seems to have effectively promoted community participation but in an uneven way, such that a clear distinction can be made between high and low participation across communities. The reasons why participation was higher in one sub-set of communities are not immediately clear, but the important finding is that, essentially, we only find statistically significant project impacts among the high-participation treatment groups. It validates the project's premise of the importance of the chosen participatory approach but leaves open the question why this was not sufficiently pursued or did not work equally well across all beneficiary watershed communities.

We can say, though, that the greater community participation in soil and water conservation practices among the high-participation sub-group was a result of the project, as perceived by those beneficiaries.[9]

(b) Beneficiaries with high participation have significantly higher incomes (17.8%) than the control group. It is unclear, however, which project activities have contributed, and how, to this positive impact. In comparison with the control group, it is not associated with better crop yields, greater income diversification or off-farm income opportunities, and neither with enhanced women's empowerment, nor reduced conflict over land. We do find significantly higher cow milk productivity and greater herd size among beneficiaries with high participation, as well as benefits from "cut-and-carry" forage collection. To a limited extent, these outcomes could partially explain the impact on incomes. The lack of impact on crop productivity or income diversification suggest that the promotion of SWC practices and income-generating activities induced no direct economic gains to beneficiary households. Part of this outcome might be explained by the fact that SWC were mostly promoted for off-farm, community resource protection (IFAD 2021), hence not directly impacting on farm productivity or household-level economic opportunities.

(c) Dietary diversity is greater among beneficiaries with high levels of participation. This outcome may be associated with the higher incomes of those beneficiaries, as well as with higher productivity of milking cows. Again, we cannot say for sure based on the evidence obtained and the finding is not corroborated by commensurate improvements in food security, as measured through the food insecurity experience scale.

(d) The analysis of geo-spatial data showed improvement in vegetation coverage over the 7-year period of observation (2013-2019) and of most of the project's period of implementation. This greening of the watersheds over time could be associated with improved erosion techniques or common land rehabilitation. However, such improvements were observed for all watersheds in the area and no statistical differences could be detected between the CBINReMP beneficiary watersheds and the control group. The potential reasons could be that such improvements may have taken place through different means in all watersheds as well as because of exogenous factors, such as the increased rainfall experienced in the LTW area during the final years of the project's implementation.

[9]This could sound as tautological, but in making the distinction between high and low level of participation we only assess ex-post levels of participation. The related variables regarding "perceived" increase in participation in SWC practices at the household and community level (see Table10) refers to comparisons of dwellers and community leaders with the situation of ten years prior to the survey data collection (March 2020), hence around the beginning stages of the project.

In summary, this ex-post assessment indicates that the CBINReMP had only very limited, quantitatively verifiable impact on rural livelihoods. It seemed to have contributed to higher household incomes and some greater dietary diversity, but only where the project managed greater community participation. However, even for those beneficiaries, livelihood conditions had not become significantly more productive, diversified, resilient, or sustainable than those of the comparison group.

Caveats and limitations of impact assessment

The lack of a proper baseline survey, incomplete information about targeted watershed communities and often lack of clear distinction lines between the project's interventions and support provided to communities through other mechanisms make it sheer impossible to identify the true impact of the CBINReMP. Four additional challenges had to be faced, including possible selection biases because of non-random placement (targeting) of the project, self-selection of beneficiaries into receiving the project, possible spatial spill-over effects of project benefits to non-treatment communities, and the project's phased rollout.

The last challenge could not be address having only an after-the-project survey to undertake the impact assessment. As discussed in greater detail in Section 3, the other three challenges could be addressed to a large extent. First, to account for the non-random placement of the project, we control for observable community-level characteristics and geographical attributes that are exogenous to the project. Admittedly, of the many possible unobservable confounders not all could be accounted for. Second, we had to consider all households living within the targeted watersheds as beneficiaries, so the findings as presented should be considered as "intent-to-treat" effects. Third, since project interventions were planned at the kebele level, they could have benefited both targeted and non-targeted watersheds within a treated kebele. We checked for the potential spatial "spillover" effect due to the kebele level planning, as described in Section 3. We did not find any systematic pattern that could point at significant "spillover" effects owing to the project's design.

This said, the findings are limited to what we were able to test with the limitations of an after-the-intervention survey and no comparison of rural livelihood and watershed management conditions at the start or mid-stage of the project's implementation. That is, we could only make an ex-post comparison between the treatment group and the comparison group, and, what is more, given apparent uneven implementation of the project's intervention across the target group we distinguished between communities with high and low participation levels. This distinction did allow us to detect discernable impacts in terms of some of the key outcome variables, suggesting that high community participation is a pre-condition for obtaining the observed impacts. In doing so, however, we could no longer assess, rigorously, whether such higher community participation was effectively induced by the project (for obvious endogeneity issues).

Recommendations for better embedding impact assessment elements in future project design

There are a few basic principles related to project design that can help better prepare projects for *ex post* impact assessment in the future. For the purpose of this discussion, we assume random assignment of projects to a portion of eligible communities is not possible. From an evaluation perspective, conducting a first phase of a project with random assignment to learn about what works in advance of a second phase in which scale up occurs is ideal, but difficult to implement in practice due to timing issues.

The basic principles are as follows. First, it is important to better track where projects will be implemented, where they will not be, and reasons that decisions are made about where to implement projects and where not to do so. Second, it is important to track what types of investments or interventions take place in specific places. Third, it is extremely helpful to collect baseline information in both communities that are to be targeted for interventions, and in similar communities that are not to be targeted. None of these principles are perhaps surprising, but it is worth emphasizing them here.

Information about targeting long-term projects is really crucial in developing ex post impact evaluations. Here, the more information that can be recorded, the better. If the analyst can know why specific places are targeted whereas others are not, they can control for those differences in analysis, and the unobservable component of potential program placement bias becomes minimized. Of course, with large programs it could become more difficult to find valid comparison groups; meaning, it could become more difficult to find comparison groups that are similar to the participant groups. However, in a situation like the CBINReMP, the ex-post evaluation was forced to make uncomfortable assumptions about these unobserved differences, in part because the watersheds in which implementation occurred were only partly known, and as discussed below the borders of those watersheds remain unknown.

A second somewhat basic change that can help *ex post* impact evaluation of complex projects like CBINReMP is to track which interventions take place in which places (in this case, in which watersheds). Complex interventions often have lots of different parts and naturally evaluators (or planners of future projects) would like to know which components are more likely to have impacts on outcomes of interest, which components might need a redesign, and which components work well together (or do not). However, answering all these questions in a convincing manner requires a substantial amount of data collection. More basically, though, the questions cannot be answered well if there are no records (or only piecemeal records) of what interventions took place where. In this case, the survey asked about it, but there was little useful variation, making questions about project components nearly impossible to answer.

Third, when projects are not randomized, having baseline data becomes quite essential. Ideally, the baseline data collection can then be used later in efforts to match participants or participant communities with like members of the control group. There are two reasons such data are essential to have, and one programming implication. A first reason is that they can help build the case that changes occur in the participant group relative to the control group in a way that cannot be done in simple before-after comparisons (or after project comparisons). Before-after comparisons are notoriously inadequate at demonstrating changes. If comparisons are made just after projects end, one has to make uncomfortable assumptions about the types of variables that would not have changed or have to use recall data that are subject to well-known errors such as telescoping. In this report, such assumptions had to be made. Programmatically, then, it is important to make decisions about some places in which projects will be implemented in advance of that baseline survey (and starting any activities), so the baseline survey is not wasted on places in which the project never occurs (this issue was an issue in RIMS surveys collected in the past).

Recommendations regarding collection of geo-spatial data for evaluation of SLM-type projects

Sections 3 and 4 indicated at the limited use this evaluation could make of geo-spatial data. Here, we provide some recommendations for a methodology that would overcome such limitations in the use of geo-spatial data for future impact evaluations of this kind.

The single most enhancing component would be an accurate depiction of watershed boundaries, without which a significant amount of error in measurement should be expected. Delineation might take two forms, one through the digitization of existing physical watershed boundary maps, and the second, through recreating the original watershed methodology used to create the watershed boundaries which would require the *exactly matching* watershed pour-point coordinates, as well the digital elevation map used to create it. Beyond these suggestions to create accurate boundaries, a handful of other improvements could be implemented. Significant effort could be made to filter out non-agricultural land from imagery at a localized level. This would enhance the measurement of the crop cycle. Additionally, all variables could be de-trended using precipitation data. Although our research indicated that the sampling design, that chose neighboring watersheds as complimentary treatment and control, makes weather patterns difficult to delineate, differences between watersheds might simply be due to the erratic nature of rainfall in Ethiopia's complex topology. Finally, a larger set of time-series features could be analyzed. For instance, looking at whether treatment watersheds are less likely to experience sudden shocks to plant health, observed by a steep decline in NDVI in any given month. While additional research could be undertaken, the basic limitation of inaccurate and approximate watershed boundaries makes these efforts less productive.

References

Baylis, Kathy,Jordi Honey-Roses, Jan Borner, Esteve Corbera, Driss Ezzine-de-Blas, Paul J. Ferraro, Renaud Lapeyre, U. Martin Persson, Alex Pfaff, and Sven Wunder 2016. "Mainstreaming impact evaluation in nature conservation." *Conservation Letters* 9 (1): 58-64. (http://doi: 10.1111/conl.12180).

Deininger, Klaus, Daniel Ayalew Ali, and Tekie Alemu. 2011. "Impacts of land certification on tenure security, investment, and land market participation: evidence from Ethiopia." *Land Economics* 87 (2): 312-334.

Deininger, Klaus, and Fang Xia. 2016. "Quantifying spillover effects from large land-based investment: the case of Mozambique." *World Development* 87: 227-241.

Deininger, Klaus, and Songqing Jin. 2006. "Tenure security and land-related investment: Evidence from Ethiopia." *European Economic Review* 50 (5): 1245-1277.

Descheemaeker, Katrien, Everisto Mapedza, Tilahun Amede, and Wagnew Ayalneh. 2010. "Effects of Integrated Watershed Management on Livestock Water Productivity in Water Scarce Areas in Ethiopia." *Physics and Chemistry of the Earth, Parts A/B/C* 35 (13–14): 723–29. (https://doi.org/10.1016/j.pce.2010.06.006)

Franklin, J.M., W. Eddings, R.J. Glynn, and S. Schneeweiss 2015. "Regularized Regression Versus the High-Dimensional Propensity Score for Confounding Adjustment in Secondary Database Analyses," *American Journal of Epidemiology* 182(7): 651-659.

Funk, Chris, Pete J. Peterson, Martin F. Landsfeld, Diego H. Pedreros, James P. Verdin, James D. Rowland, Bo E. Romero, Gregory J. Husak, Joel C. Michaelsen, and Andrew P. Verdin. 2014. "A Quasi-Global Precipitation Time Series for Drought Monitoring." *United States Geological Survey Data Series* 832 (4): 1–12.

Gebregziabher Gebrehaweria, Dereje Assefa Abera, Girmay Gebresamuel, Meredith Giordano and Simon Langan. 2015. An Assessment of Integrated Watershed Management in Ethiopia. IWMI Working paper No. 170. Colombo: International Water Management Institute. (https://www.iwmi.cgiar.org/publications/iwmi-working-papers/iwmi-working-paper-170/)

Gebreselassie, Samuel, Oliver K. Kirui, and Alisher Mirzabaev. 2016. "Economics of land degradation and improvement in Ethiopia." In: *Economics of land degradation and improvement–a global assessment for sustainable development*. Nkonya, Ephraim, Alisher Mirzabaev, and Joachim von Braun (Eds.). Springer Open, Cham, pp.401-430.

Holden, S.T., Deininger, K., and Ghebru, H. 2009. 'Impacts of low-cost land certification on investment and productivity', *American Journal of Agricultural Economics* 91(2): 359–373.

IFAD 2009a. Proposed loan and grant to the Federal Democratic Republic of Ethiopia for the Community-based Integrated Natural Resources Management Project,

____ 2009b. Project Design Report.

____ 2009c. Project Implementation Manual,

____ 2010-18. Project Supervision Reports.

____ 2014. Mid-term Review Report.

____ 2016. Guidelines for the Project Completion Report Validations and Project Performance Assessments.

____ 2019. Project Completion Report.

____ 2021. *Ethiopia: Community-based Integrated Natural Resources Management Project: Impact Evaluation*. Rome: International Fund for Agricultural Development. (https://www.ifad.org/documents/38714182/43698779/ETHIOPIA+IE+%2B+COVER.pdf/a2abdf7c-093b-17a2-7c67-572663d21baf?t=1631619623047)

Kassie, Menale, John Pender, Mahmud Yesuf, Gunnar Köhlin, Randy Bluffstone, and Elias Mulugita. 2008. Estimating returns to soil conservation adoption in the northern Ethiopian highlands. *Agricultural Economics* 38 (2): 213-232. (https://doi.org/10.1111/j.1574-0862.2008.00295.x)

Kassie, Menale, Stein Holden, Gunnar Köhlin, and Randy Bluffstone. 2009. Economics of soil conservation adoption in high-rainfall areas of the Ethiopian Highlands. University of Göttingen Working Paper 400, Göttingen. (https://www.researchgate.net/publication/46470445_Economics_of_Soil_Conservation_Adoption_in_High-Rainfall_Areas_of_the_Ethiopian_Highlands)

Kato, Edward, Dawit Mekonnen, Solomon Tiruneh, Claudia Ringler. 2019. Sustainable land management and its effects on water security and poverty: Evidence from a watershed intervention program in Ethiopia. *IFPRI Discussion Paper* 1811. Washington D.C.: International Food Policy Research Institute (http://ebrary.ifpri.org/utils/getfile/collection/p15738coll2/id/133144/filename/133355.pdf)

Mengistu, Fekadu, and Engdawork Assefa. 2019. "Enhancing Livelihood Assets of Households through Watershed Management Intervention Program: Case of Upper Gibe Basin, Southwest Ethiopia." *Environment, Development and Sustainability*, November. (https://doi.org/10.1007/s10668-019-00534-x)

Nkonya, Ephraim, Alisher Mirzabaev, and Joachim von Braun. 2016. "Economics of land degradation and improvement: an introduction and overview." In: *Economics of land degradation and improvement–a global assessment for sustainable development*. Nkonya, Ephraim, Alisher Mirzabaev, and Joachim von Braun (Eds.). Springer Open, Cham, pp. 1-14.

Ostvald, Madeleine and Sabine Henders. 2014. Accounting methods for international land-related leakage and distant deforestation drivers. *Ecological Economics*, **99**:21–28. (https://doi.org/10.1016/j.ecolecon.2014.01.005)

Pender, John; and Gebremedhin, Berhanu. 2006. Land Management, Crop Production, and Household Income in the Highlands of Tigray, Northern Ethiopia: An Econometric Analysis. In: *Strategies for sustainable land management in the East African Highlands*. Pender, John, Place, Frank, and Ehui, Simeon K. (Eds.) Chapter 5. 107-140. Washington, D.C.: International Food Policy Research Institute (IFPRI). (http://ebrary.ifpri.org/cdm/ref/collection/p15738coll2/id/129592)

Słoczyński, Tymon, and Jeffrey M. Wooldridge. 2018, "A general double robustness result for estimating average treatment effects." *Econometric Theory* 34 (1): 112-133.

Schmidt, Emily, and Debay Tadesse. 2012. Household and plot level impact of sustainable land and watershed management (SLWM) practices in the Blue Nile. IFPRI ESSP Working Paper 42. Addis Ababa: International Food Policy Research Institute. (http://ebrary.ifpri.org/utils/getfile/collection/p15738coll2/id/127070/filename/127281.pdf)

Siraw, Zewdu, Woldeamlak Bewketand Mekonnen Adnew Degefu. 2018. Assessment of livelihood benefits of community-based watershed development in northwestern highlands of Ethiopia, *International Journal of River Basin Management.* (https://doi.org.10.1080/15715124.2018.1505733)

Swindale, Anne and Paula Bilinsky. 2006. Household Dietary Diversity Score (HDDS) for Measurement of Household Food Access: Indicator Guide VERSION 2 September (https://www.fantaproject.org/sites/default/files/resources/HDDS_v2_Sep06_0.pdf)

Tibishirani, R. J. 1996. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B*58: 267–288.

Wooldridge, Jeffrey M. 2007. "Inverse probability weighted estimation for general missing data problems." *Journal of Econometrics* 141 (2): 1281-1301.

APPENDIX Supplementary Results Tables

Table A.1 Average treatment effect by control subgroups: "spillover" effect

Outcome Variable	Treated (N = 887) [A]	Control_T[a] (N = 240) [B]	Control_C (N = 493) [C]	Wald test (F value) [A=B]	Wald test (F value) [A=C]
Lead outcome variables					
Food security	7.62	7.08	7.59	**11.96*** **	0.03
Dietary diversity	2.14	2.03	2.31	0.10	0.33
Total income (log)	9.31	9.41	9.23	0.55	0.44
Asset holding	2.89	3.44	2.68	**6.59** **	1.66
Social cohesion	0.01	-0.27	0.11	1.51	0.32
Participation in WS plan	0.68	0.68	0.69	0.01	0.08
Membership in grazing land	0.51	0.46	0.49	0.67	0.20
White teff yield	1.52	1.52	1.62	0.00	1.01
Black teff yield	1.67	1.60	1.59	0.17	0.37
Maize yield	3.07	3.17	3.03	0.45	0.12
Lactation period	2.03	1.94	2.05	**3.46***	0.22
Cow productivity	0.11	0.05	0.09	1.85	0.58
Fattening period	1.22	1.29	1.30	0.33	0.37
Intermediate outcome variables					
Income diversification	1.59	1.65	1.57	1.34	0.18
Free grazing	2.74	3.19	2.85	2.43	0.10
Female WS committee	12.26	12.73	11.63	0.03	0.05
Resilience to climate change	0.77	0.83	0.65	0.36	1.35
Agri. productivity (10 years)	0.51	0.49	0.42	0.03	0.76
Off-farm income availability	0.83	0.93	0.61	1.34	**5.11** **
SWC communal land (10 years)	0.49	0.46	0.44	0.24	0.88
Labour time for terracing	102.73	83.06	81.82	0.58	0.86
Labour time for cut off drainage	59.8	26.32	39.37	4.11**	1.35
Labour time for gully rehabilitation	38.24	21.95	50.15	3.70*	0.67
Labour time for tree planting	275.5	24.52	30.25	1.12	1.07
SWC own land (10 years)	0.51	0.45	0.47	1.62	0.61
Cereal yield (10 years)	0.52	0.49	0.45	0.34	1.96
Herd size (10 years)	0.41	0.40	0.31	0.03	**8.72*** **
Cut and carry	0.77	0.77	0.74	0.02	1.13
Resource conflict	0.24	0.19	0.25	1.41	0.11

Note: [a] Control_T and Control_C refer to control groups located within and outside of the treated kebeles, respectively.

Table A.2 Balancing test of community level co-variates using propensity score weights [N = 134]

No.	Variable code	Measurement and definition	Treated	Control	Wald test (prob > F)
1	si_i18c	Severity of conflicts before 10 years (Number of conflicts/year)	12.02	15.63	0.31
2	sc_c32b	Percentage of HH who adopt energy efficient tech. before 10 years	6.20	7.20	0.79
3	kolla_agro	=1 if kola covers >25percent and 0 otherwise	0.09	0.08	0.74
4	woyandega_agro	=1 if midland covers >25 percent and 0 otherwise	0.90	0.87	0.58
5	dega_agro	=1 if lowland covers >25 percent and 0 otherwise	0.09	0.08	0.71
6	ws_distance	Distance from community watershed to the center of the kebele (km)	2.80	2.68	0.78
7	vehicle_access	=1 if there is vehicle access and 0 otherwise	0.16	0.15	0.83
8	truck_access	=1 if there is truck access and 0 otherwise	0.02	0.06	0.31
9	wscomm_09	=1 if watershed committee is formed before 2009 and 0 otherwise	0.95	0.93	0.43
10	sc_c26c1	Area of forest rehabilitated before 10 years (ha)	8.56	8.88	0.92
11	sc_c5a2	Area of land allocated for crop production before 10years (ha)	299.1	280.4	0.54
12	sc_c5b2	Area of land allocated for pasture/grazing before 10years (ha)	60.09	66.71	0.48
13	sc_c5c2	Area of land allocated for forest before 10years (ha)	52.5	63.66	0.33
14	sc_c5d2	Area of land covered by degraded land before 10years (ha)	8.59	10.30	0.41
15	sd_d6a1	Local cow productivity in 2009 (liters of milk/cow/day)	1.97	1.85	0.56
16	sd_d6b1	Local chicken productivity in 2009 (egg/hen/year)	98.25	109.58	0.34
17	sd_d6c1	Honey productivity from traditional beehive in 2009 (kg/hive/year)	11.50	12.17	0.52
18	sd_d6d1	Fattening period of cattle in 2009 (months)	4.93	3.83	0.51

Note: The results on the wald test column are p values.

Table A.3 Participation variables used to redefine treatment status

No.	Activities (planned to be) implemented by the project	Variable definition and measurement	Remark
A	***Participatory watershed management***		
1	o Participation in watershed management plan	= 1 if the HH participate in the community level watershed plan	HH[H4]
2		= 1 if there is community level watershed management plan	C [C15]
3		= 1 if the community watershed management plan was participatory	C [C16]
4		= 1 if there is kebele level watershed management plan	C [C18]
5		= 1 if the kebele watershed management plan was participatory	C [C19]
B	***Pasture and forage development***		
6	o Community bylaws	= 1 if there is written by law to administer watershed	C [C14]
7	o Free grazing and grazing land associations	= 1 if the HH is a member of grazing land association	HH [H16]
8		= 1 if the HH practices cut and carry or controlled grazing	HH [H17]
9		= 1 if the HH practices free grazing	HH [G13]
C	***Soil and water conservation***		
10	o Participation in community level SWC practices	= 1 if SWC practices are implemented [plot level data]	HH[D10a]
11		= 1 if the HH participate on terrace construction	HH [H19a]
12		= 1 if the HH participate on cutoff drain	HH [H19b]
13		= 1 if the HH participate on gully rehabilitation	HH [H19c]
14		= 1 if the HH participate on tree planting	HH [H19d]
15		= 1 if the HH participate on area closure	HH [H19e]
16		= 1 if the HH participate on forage development	HH [H19f]
17	o Training on SWC	= 1 if the HH got training on soil and water conservation	HH [H8]
D	***Biodiversity and ecosystem***		
18	o Meeting/training on biodiversity	= 1 if there was consultative meeting/training on biodiversity	C [F2]

Note: HH and C in the remark column refer to household and community questionnaires, respectively.

Table A.4Definition and measurement of outcome variables

Outcome Variables	Definition and measurement
Lead outcome variables	
Dietary diversity	HH dietary diversity score estimated using the 12 standard food groups listed on section L of household questionnaire (the score ranges from zero to 12)
Food security	Experience based food security index: generated from recall of the typical week consumption of the household (refer M1 of HH questionnaire for details)
Total income	Log: Total income from crop, livestock, on/off-farm sources in Ethiopian birr
Asset holding	Constructed wealth category ranging from 1st to 5th quantiles where 1st refers to the poorest and 5th is the richest.
Social cohesion	Social cohesion index: computed interitem correlation of the questions on H2 of the community questionnaire
White teff yield	Household level average productivity of white teff [quintal/hectare ; in log]
Black teff yield	Household level average productivity of black teff [quintal/hectare; in log]
Maize yield	Household level average productivity of maize [quintal/hectare; in log]
Lactation period	Household level average lactation period of local cow [months; in log]
Cow milk productivity	Household level average productivity of local cow [milk/cow/day; in log]
Fattening period	Household level average fattening period of sheep/goat [months; in log]
Intermediate outcome variables	
Cut and carry	= 1if the household practice cut and carry system and 0 otherwise
Free grazing	Area of land allocated for free grazing [ha; in log]
Resource conflict	= 1 if the household involved in land related disputes and 0 otherwise
Labor for community works	Total labor time allocated for community works (i.e. - terrace, cut off, tree planting, and gully rehabilitation) [labour time in hours/year]
Resilience to climate change	= 1 if community coping capacity has improved compared to 10 years ago
Income diversification	Number of income sources of the household
Off farm income	= 1if availability of off/on farm income increased compared to 10 years ago
Female WS committee	Female watershed committee members [percent]
Agri. productivity (10 yr)	= 1if productivity of cereal or livestock increased compared to 10 years ago
Cereal yield (10 yr)	= 1if productivity of cereal increased compared to 10 years ago
Herd size (10 years)	= 1if herd size increased compared to 10 years ago
SWC on own land (10 yr)	= 1if participation in own land SWC increased compared to 10 years ago
SWC on common land (10 yr)	= 1if participation in communal land SWC increased compared to 10 years ago
Participation in WS plan	= 1if the household participate in community level watershed plan, 0 otherwise
Membership in grazing land	= 1if the household a member of grazing land association, 0 otherwise

Table A.5a Descriptive statistics: Lead outcome variables

Outcome Variable	Definition and measurement of variables	Control HH/watershed				Treated household/watershed				All households/watersheds				Remark
		N	Mean	SD	Skewness	N	Mean	SD	Skewness	N	Mean	SD	Skewness	
A. Socioeconomic outcomes														
Food security	HH dietary diversity score	733	7.39	1.53	0.10	887	7.63	1.63	0.17	1655	7.53	1.60	0.15	HH
	Experience based food security index	733	2.25	2.36	0.85	887	2.14	2.04	0.87	1655	2.17	2.32	0.87	HH
Total income and assets	Total income from crop, livestock, off-farm, and on-farm activities (log)	655	9.35	1.36	-0.06	800	9.31	1.41	0.04	1487	9.39	1.38	-0.01	HH
	Income diversification = Number of income sources	733	1.58	0.82	-0.14	887	1.59	0.78	-0.14	1655	1.59	0.79	-0.15	HH
	Asset holding (constructed wealth category)	733	2.91	1.41	0.05	887	2.89	1.42	0.07	1655	2.89	1.41	0.06	HH
Social cohesion	An index generated from five questions (i.e. measuring interitem correlations)	733	0.01	0.90	0.58	887	0.01	0.74	-0.32	1655	0.000	0.82	0.25	HH
B. Adaptation to climate change														
Adaptation to climate change	= 1 if the HH take adaptation measures	733	0.07	0.25	3.34	887	0.07	0.26	3.27	1655	0.07	0.25	3.34	HH
	= 1 if coping capacity of the HH has improved compared to 10 years ago	733	0.32	0.46	0.78	887	0.33	0.47	0.71	1655	0.32	0.47	0.73	HH
	= 1 if coping capacity of the community has improved compared to 10 years ago	733	0.72	0.45	-0.98	887	0.77	0.42	-1.28	1655	0.74	0.43	-1.13	HH
C. Agricultural productivity														
Crop productivity	White teff yield [qt/ha] (log)	216	1.58	0.80	-0.65	287	1.52	0.79	-0.15	519	1.54	0.80	-0.38	HH

	Black teff yield [qt/ha] (log)	151	1.54	0.68	-0.39	152	1.67	0.69	-0.10	308	1.61	0.68	-0.23	HH
	Maize yield [qt/ha] (log)	244	3.07	0.92	-0.54	263	3.07	0.79	-0.44	520	3.07	0.85	-0.49	HH
Livestock productivity	Local cow lactation period (month) (log)	578	2.00	0.40	-0.46	713	2.03	0.40	-0.40	1320	2.02	0.40	-0.44	HH
	Local cow productivity (milk/cow/day) (log)	570	0.08	0.45	0.48	704	0.11	0.49	1.97	1303	0.09	0.47	1.37	HH
	Fattening period of local sheep/goat (month) (log)	377	1.25	0.69	2.77	463	1.22	0.67	3.19	858	1.24	0.68	2.99	HH

Table A.5b Descriptive statistics: Lead outcome variables by participation level

Outcome Variable	Definition and measurement of variables	Control HH/watershed				High-participation Treated household/watershed				Low-participation Treated household/watershed				Remark
		N	Mean	SD	Skewness	N	Mean	SD	Skewness	N	Mean	SD	Skewness	
A. Socioeconomic outcomes														
Food security	HH dietary diversity score	733	7.39	1.53	0.10	524	7.81	1.59	0.16	363	7.36	1.66	0.25	HH
	Experience based food security index	733	2.25	2.36	0.85	524	2.18	2.34	0.88	363	2.08	2.25	0.83	HH
Total income and assets	Total income from crop, livestock, off-farm, and on-farm activities (log)	655	9.35	1.36	-0.06	478	9.43	1.40	0.30	322	9.14	1.42	-0.31	HH
	Income diversification = Number of income sources	733	1.58	0.82	-0.14	524	1.67	0.78	-0.12	363	1.48	0.75	-0.23	HH
	Asset holding (constructed wealth category)	733	2.91	1.41	0.05	524	3.01	1.42	-0.03	363	2.72	1.40	0.21	HH
Social cohesion	An index generated from five questions (i.e. measuring interitem correlations)	733	0.01	0.90	0.58	524	-0.04	0.76	-0.14	363	0.09	0.72	-0.59	HH
B. Adaptation to climate change														
Adaptation to climate change	= 1 if the HH take adaptation measures	733	0.07	0.25	3.34	524	0.10	0.29	2.68	363	0.04	0.18	4.99	HH
	= 1 if coping capacity of the HH has improved compared to 10 years ago	733	0.32	0.46	0.78	524	0.38	0.48	0.47	363	0.25	0.43	1.13	HH
	= 1 if coping capacity of the community has improved compared to 10 years ago	733	0.72	0.45	-0.98	524	0.81	0.39	-1.61	363	0.71	0.45	-0.91	HH
C. Agricultural productivity														
Crop productivity	White teff yield [qt/ha] (log)	216	1.58	0.80	-0.65	177	1.57	0.82	0.03	110	1.45	0.76	-0.58	HH

	Black teff yield [qt/ha] (log)	151	1.54	0.68	-0.39	101	1.70	0.72	-0.27	51	1.61	0.63	0.32	HH
	Maize yield [qt/ha] (log)	244	3.07	0.92	-0.54	122	3.04	0.83	-0.40	141	3.11	0.76	-0.47	HH
Livestock productivity	Local cow lactation period (month) (log)	578	2.00	0.40	-0.46	435	2.03	0.40	-0.29	278	2.02	0.41	-0.56	HH
	Local cow productivity (milk/cow/day) (log)	570	0.08	0.45	0.48	430	0.15	0.46	0.31	274	0.05	0.53	3.84	HH
	Fattening period of local sheep/goat (month) (log)	377	1.25	0.69	2.77	295	1.22	0.69	3.11	168	1.23	0.64	3.35	HH

Table A.6a Descriptive statistics: Intermediate outcome variables

Outcome Variable	Definition and measurement of variables	Control HH/watershed				Treated household/watershed				All households/watersheds				Remark
		N	Mean	SD	Skewness	N	Mean	SD	Skewness	N	Mean	SD	Skewness	
A. Land certification: resource allocation, credit access and woman empowerment														
Resource allocation and decision making	= 1 if a woman is holder of land certificate	733	0.06	0.24	3.51	887	0.07	0.25	3.37	1655	0.06	0.25	3.48	HH
	= 1 if land certificate improves position of a woman	733	0.94	0.23	-3.75	887	0.95	0.22	-4.04	1655	0.95	0.22	-3.95	HH
	= 1 if the wife is responsible to sell the crop	726	0.60	0.49	-0.41	874	0.65	0.47	-0.64	1635	0.63	0.48	-0.54	HH
	Female watershed committee members [Percent]	709	15.09	14.8	0.62	887	12.23	13.39	0.80	1631	13.72	14.2	0.69	Comm.
Land investment	= 1 if the HH undertakes long-term SWC practices	733	0.62	0.48	-0.48	887	0.65	0.47	-0.66	1655	0.64	0.47	-0.60	HH
Credit access	= 1 if HH believes land certificate improves access to credit	681	0.94	0.22	-3.39	807	0.96	0.18	-5.08	1521	0.95	0.20	-4.52	HH
Resource conflict	= 1 if the HH encounter land related disputes	733	0.23	0.42	1.27	887	0.24	0.43	1.18	1655	0.24	0.42	1.22	HH
	= 1 if the HH encounter water or forest related disputes	733	0.09	0.29	2.75	887	0.13	0.33	2.26	1655	0.11	0.31	2.46	HH
B. Natural resource management														
Soil and water conservation	Stone/soil bund/stone faced soil bund (meter)	565	12921	18787	1.49	816	20686.6	47956.5	3.89	1416	17130	38538	4.60	Comm.
	Cut off drain (meter)	505	2302.6	5266	3.86	575	4494.7	15243.8	5.40	1103	3416.1	11620	6.73	Comm.
	Gully rehabilitation (meter)	398	5854	16128.9	3.88	682	825.9	1677.8	3.54	1092	2707.3	10115	6.60	Comm.
	Tree planting (number)	505	28.12	54.12	3.13	743	52.64	252.5	7.53	1283	191.2	1549	9.96	Comm.
Labor spent on community level	Labor hour spend on terrace construction [labor hour/yr]	444	84.47	168.28	6.34	549	102.73	339.93	9.50	1016	93.58	273.6	10.53	HH

conservation practices	Labor hour spend on cut off drain [labor hour/yr]	148	36.70	101.39	8.45	190	59.98	197.47	8.34	351	48.52	159.8	9.52	HH
	Labor hour spend on gully rehabilitation [labor hour/yr]	155	42.08	117.59	6.92	209	38.24	91.43	10.17	378	39.23	101.4	8.53	HH
	Labor hour spend on tree planting [labor hour/yr]	155	32.91	96.75	8.77	208	275.51	3465.15	14.31	373	167.53	2588	19.21	HH
Flooding	= 1 if the HH experienced high flooding	214	0.48	0.50	0.05	290	0.55	0.49	-0.19	520	0.52	0.49	-0.09	HH
	= 1 if flooding is more severe compared to 10 years ago	733	0.27	0.44	1.00	887	0.28	0.45	0.93	1655	0.28	0.45	0.94	HH
Free grazing	Area of land allocated for free grazing [Hectare]	733	36.24	60.04	3.15	887	21.67	28.45	2.52	1655	28.36	45.64	3.81	Comm.
Nursery access	Distance to the nearest nursery site from home [minutes]	733	21.54	102.3	7.10	887	21.81	69.8	-0.24	1655	22.47	86.27	5.16	Comm.
	Distance to the nearest nursery site from the center of the community [minutes]	733	45.04	50.0	0.86	887	78.98	346.9	8.09	1655	62.71	256.7	10.87	Comm.
Water flow	= 1 if the flow of river and springs has reduced	733	0.43	0.49	0.27	887	0.45	0.49	0.19	1655	0.44	0.49	0.23	HH
C. Water harvesting and energy efficient technologies														
Water harvesting	= 1 if the HH adopted water harvesting technology	733	0.05	0.22	4.10	887	0.45	0.49	0.19	1655	0.05	0.22	4.12	HH
Access to energy efficient technologies	=1 if the HH adopted the technology and 0 otherwise	733	0.09	0.29	2.77	887	0.12	0.32	2.38	1655	0.11	0.31	2.57	HH
	Households in the community who adopt energy efficient technology [percent]	326	23.89	28.53	1.21	516	35.32	35.04	0.49	842	30.89	33.13	0.74	Comm.

Note: HH and Comm. refer to household and community level data, respectively.

Table A.6b Descriptive statistics: Intermediate outcome variables by participation level

Outcome Variable	Definition and measurement of variables	Control HH/watershed				High-participation Treated household/watershed				Low-participation Treated household/watershed				Remark
		N	Mean	SD	Skewness	N	Mean	SD	Skewness	N	Mean	SD	Skewness	
A. Land certification: resource allocation, credit access and woman empowerment														
Resource allocation and decision making	= 1 if a woman is holder of land certificate	733	0.06	0.24	3.51	524	0.06	0.24	3.59	363	0.08	0.27	3.09	HH
	= 1 if land certificate improves position of a woman	733	0.94	0.23	-3.75	524	0.94	0.23	-3.89	363	0.95	0.21	-4.29	HH
	= 1 if the wife is responsible to sell the crop	726	0.60	0.49	-0.41	517	0.66	0.47	-0.68	357	0.64	0.48	-0.59	HH
	Female watershed committee members [Percent]	709	15.09	14.8	0.62	524	12.57	13.22	0.76	363	11.71	13.63	0.87	Comm.
Land investment	= 1 if the HH undertakes long-term SWC practices	733	0.62	0.48	-0.48	524	0.76	0.42	-1.25	363	0.50	0.50	-0.01	HH
Credit access	= 1 if HH believes land certificate improves access to credit	681	0.94	0.22	-3.39	485	0.96	0.19	-4.75	322	0.97	0.16	-5.73	HH
Resource conflict	= 1 if the HH encounter land related disputes	733	0.23	0.42	1.27	524	0.25	0.43	1.14	363	0.23	0.42	1.25	HH
	= 1 if the HH encounter water or forest related disputes	733	0.09	0.29	2.75	524	0.13	0.33	2.20	363	0.11	0.32	2.36	HH
B. Natural resource management														
Soil and water conservation	Stone/soil bund/stone faced soil bund (meter)	565	12921	18787	1.49	475	19884	46683	4.23	341	21803	49722	3.48	Comm.
	Cut off drain (meter)	505	2302.6	5266	3.86	337	4627	15138	5.44	238	4306	15421	5.35	Comm.
	Gully rehabilitation (meter)	398	5854	16128.9	3.88	428	942.74	1807.2	3.19	254	629	1415	4.40	Comm.
	Tree planting (number)	505	28.12	54.12	3.13	442	73.42	324.8	5.76	301	22.12	31.73	2.00	Comm.
Labor spent on community level	Labor hour spend on terrace construction [labor hour/yr]	444	84.47	168.28	6.34	429	108.32	342.9	9.32	120	82.76	329	10.27	HH

conservation practices	Labor hour spend on cut off drain [labor hour/yr]	148	36.70	101.39	8.45	190	62.48	202.54	8.12	10	15.05	16.4	2.31	HH
	Labor hour spend on gully rehabilitation [labor hour/yr]	155	42.08	117.59	6.92	192	39.06	95.12	9.81	17	29.05	23.94	1.11	HH
	Labor hour spend on tree planting [labor hour/yr]	155	32.91	96.75	8.77	194	293.4	3587.9	13.81	14	26.85	33.96	1.25	HH
Flooding	= 1 if the HH experienced high flooding	214	0.48	0.50	0.05	185	0.55	0.49	-0.23	105	0.53	0.50	-0.13	HH
	= 1 if flooding is more severe compared to 10 years ago	733	0.27	0.44	1.00	524	0.27	0.45	1.01	363	0.31	0.46	0.84	HH
Free grazing	Area of land allocated for free grazing [Hectare]	733	36.24	60.04	3.15	524	21.71	29.23	2.46	363	21.50	27.3	2.62	Comm.
Nursery access	Distance to the nearest nursery site from home [minutes]	733	21.54	102.3	7.10	524	31.18	65.01	-0.17	363	8.27	74.17	-0.18	Comm.
	Distance to the nearest nursery site from the center of the community [minutes]	733	45.04	50.0	0.86	524	28.94	38.29	2.56	363	151.2	532.5	5.07	Comm.
Water flow	= 1 if the flow of river and springs has reduced	733	0.43	0.49	0.27	524	0.42	0.49	0.34	363	0.50	0.50	-0.02	HH

C. Water harvesting and energy efficient technologies

Water harvesting	= 1 if the HH adopted water harvesting technology	733	0.05	0.22	4.10	524	0.07	0.25	3.47	363	0.02	0.13	6.99	HH
Access to energy efficient technologies	=1 if the HH adopted energy efficient technology and 0 otherwise	733	0.09	0.29	2.77	524	0.15	0.35	1.97	363	0.07	0.26	3.32	HH
	Households in the community who adopt energy efficient technology [percent]	326	23.89	28.53	1.21	340	38.97	34.76	0.29	176	28.26	34.58	0.93	Comm.

Table A.7 Descriptive statistics: Project participation variables

No.	Activities of the project	Variable definition and measurement	Control HH/watershed				Treated household/watershed				All households/watersheds				Remark
			N	Mean	SD	Skewness	N	Mean	SD	Skewness	N	Mean	SD	Skewness	
A	***Participatory watershed management***														
1	Participation in watershed management plan	= 1 if the HH participate in the community WS plan	733	0.68	0.46	-0.77	887	0.68	0.46	-0.80	1655	0.68	0.46	-0.80	HH
2		= 1 if there is community level WS management plan	733	0.77	0.42	-1.29	887	0.86	0.34	-2.14	1655	0.82	0.37	-1.72	Comm.
3		= 1 if the community WS management plan was participatory	733	0.90	0.29	-2.69	887	0.93	0.25	-3.44	1655	0.92	0.27	-3.10	Comm.
4		= 1 if there is kebele level WS management plan	733	0.93	0.24	-3.51	887	0.94	0.22	-3.94	1655	0.94	0.23	-3.78	Comm.
5		= 1 if kebele WS management plan was participatory	733	0.90	0.29	-2.69	887	0.93	0.25	-3.44	1655	0.92	0.27	-3.10	Comm.
B	***Pasture and forage development***														
6	Community bylaws, free grazing and grazing land associations	= 1 if there are written by-laws to administer watershed	773	0.87	0.3	-2.19	887	0.89	0.31	-2.52	1655	0.88	0.32	-2.39	Comm.
7		= 1 if the HH is a member of grazing land association	733	0.46	0.49	0.12	887	0.51	0.50	-0.05	1655	0.49	0.50	0.02	HH
8		= 1 if the HH practices cut and carry or controlled grazing	733	0.75	0.42	-01.19	887	0.78	0.41	-1.34	1655	0.77	0.42	-1.28	HH
9		= 1 if the HH practices free grazing	733	0.67	0.47	-0.72	887	0.65	0.47	-0.64	1655	0.66	0.47	-0.67	HH
C	***Soil and water conservation***														
10	Participation in community level SWC practices	= 1 if HH implemented SWC practices at one or more plot	733	0.82	0.37	-1.74	887	0.82	0.37	-1.75	1655	0.83	0.37	-1.77	HH
11		= 1 if the HH participate on terrace construction	733	0.68	0.46	-0.77	887	0.69	0.45	-0.86	1655	0.69	0.46	-0.83	HH
12		= 1 if the HH participate on cutoff drain	733	0.23	0.42	1.24	887	0.23	0.42	1.22	1655	0.24	0.42	1.21	HH
13		= 1 if the HH participate on gully rehabilitation	733	0.23	0.42	1.24	887	0.25	0.43	1.11	1655	0.25	0.43	1.14	HH
14		= 1 if the HH participate on tree planting	733	0.24	0.42	1.19	887	0.27	0.44	1.03	1655	0.26	0.43	1.08	HH
15		= 1 if the HH participate on area closure	733	0.12	0.32	2.32	887	0.13	0.33	2.17	1655	0.13	0.33	2.19	HH

No.	Activities of the project	Variable definition and measurement	Control HH/watershed				Treated household/watershed				All households/watersheds				Remark
			N	Mean	SD	Skewness	N	Mean	SD	Skewness	N	Mean	SD	Skewness	
16		= 1 if the HH participate on forage development	733	0.07	0.25	3.42	887	0.06	0.24	3.63	1655	0.06	0.24	3.84	HH
17	Training on SWC	= 1 if the HH got training on soil and water conservation	733	0.26	0.43	1.09	887	0.28	0.45	0.93	1655	0.27	0.44	0.99	HH
D	***Biodiversity and ecosystem***														
18	Meeting/training on biodiversity	= 1 if there was consultative meeting on biodiversity	733	0.32	0.47	0.72	887	0.50	0.50	-0.002	1655	0.43	0.49	0.26	Comm.

Note: The final column marks whether the variables relate to the household (HH) or community (Comm) questionnaires.

ALL IFPRI DISCUSSION PAPERS

All discussion papers are available **here**

They can be downloaded free of charge

INTERNATIONAL FOOD POLICY RESEARCH INSTITUTE

www.ifpri.org

IFPRI HEADQUARTERS

1201 Eye Street, NW

Washington, DC 20005 USA

Tel.: +1-202-862-5600

Fax: +1-202-862-5606

Email: ifpri@cgiar.org