

Draft Approach Paper

Impact Evaluation of Rural Water Supply and Sanitation Program in Ethiopia

Content

1. Introduction.....	1
2. Background: Rural Water Supply and Sanitation Program in Ethiopia.....	2
3. IE coverage	3
4. IE questions.....	4
5. IE methodology.....	5
6. Implementation arrangements.....	16
7. Budget.....	19
Annex 1: Theory of change Ethiopia RWSSP	20
Annex 2: Key parameter values used in the estimation of sample sizes.....	21

1. Introduction

The Independent Evaluation Department (IDEV) of the African Development Bank Group (AfDB) is planning to conduct an impact evaluation (IE) of the AfDB supported Rural Water Supply and Sanitation Program (RWSSP) in Ethiopia. The objective of the IE is to provide credible estimates of effects of the rural water supply, sanitation and hygiene interventions on final target beneficiary communities and households, assess if interventions work as expected, and assess sustainability of results. The purpose is to account for the interventions and support provided and to derive lessons and recommendations to improve future impact. The impact evaluation serves as an information source for the ongoing AfDB Ethiopia Country Strategy and Program Evaluation and planned Water sector thematic evaluation. The main intended users of the IE are the AfDB Board and Senior Management, Sector Department (OWAS) and Ethiopia Field Office (FO), and the Government of Ethiopia (GoE) Ministry of Water, Irrigation and Energy (MoWIE) and Ministry of Health (MoH).

The evaluation team for the IE is composed of Rita Tesselaar (IDEV Task Manager), Girma Kumbi (IDEV co-Task Manager) and Dr. Degnet Abebaw (expert quantitative impact analysis). The evaluation team undertook a preparatory mission end of July 2014 during which stakeholders were consulted, informants at different government and local levels were interviewed, program sites were visited and the IE methodological design and implementation was prepared and discussed. The mission was followed by an inception phase during which the impact evaluation design was elaborated, survey instruments were developed, implementation arrangements were made and the work plan was updated. The IE is scheduled to be completed by April 2015.

2. Background: Rural Water Supply and Sanitation Program in Ethiopia

The RWSSP started in 2006 and was funded under the Bank Group's Rural Water Supply and Sanitation Initiative. The program was extended up to end of 2014. Ethiopia was among the five countries selected for initial implementation of this Initiative which seeks to assist Regional Member Countries achieve the Millennium Development Goals (MDGs) and the African Water Vision targets for water supply and sanitation in rural areas. The program provided a contribution to the Government of Ethiopia (GoE) National Water Supply and Sanitation Programming rural component. The AfDB is a major donor in RWSS, with the financial grant to Ethiopia equivalent to UA 43.6 mln. The program disbursed 99.9% of its funding.

In 2005 rural water supply and sanitation coverage in Ethiopia, standing at 24% and 8% respectively, was particularly low. Under the National Water Supply and Sanitation Program (NWSSP), GoE intends to increase the national coverage to 62% and 54% respectively by 2015, through application of the Demand Responsive Approach (DRA) to services delivery.

Ethiopia has a federal system with nine regional governments and two city administrations. The regional governments are further decentralized to districts (woredas). Of around 800 districts in the country, the Rural Water Supply and Sanitation Program has been implemented in 125 districts.

The objective of the program is to improve access to water supply and sanitation services and thereby contribute to the achievement of MGD related goals. The main components of the program are:

Water supply: provision of new and rehabilitation of existing water supply services.

Technology choices for water supply range from hand dug wells, drilled wells, protected springs, boreholes, gravity schemes, water harvesting, subsurface dams, small piped systems to pumping systems (hand pumps, motorized, solar or windmill pumps).

Sanitation services: provision of latrine facilities for schools, health centres, communal latrines in crowded settlement areas, and demonstration latrines for associations, cooperatives and communities, and conduct of hygiene and health education campaign;

Capacity building: community sensitization, awareness raising, organization, training in water supply and sanitation; capacity building at the MOWR, MOH, Regional Water Bureaus, Regional Health Bureaus, Woreda Water Desks and Woreda Health Desks and Woreda Support Groups; capacity building for community facilitation teams, local service providers, artisans, health extension workers and spare parts suppliers for development of supply chains;

Program support: at the federal, regional and local government level institutions, and to communities; establishment of a database, annual financial and technical audits; monitoring and evaluation; and technical assistance to the MoWIE and MoH.

Of a total of 4310 planned schemes 4159 (96.5%) been constructed up to end of 2013. An estimated 1,900,000 people have been provided with access to safe water supply and sanitation. Community awareness has been created on the importance of hygiene and sanitation services.

Key outcome variables of interest are: access to and use of improved water sources and toilet facilities; quality of water; water consumption for domestic use; hygiene practices; travel distance to improved water source; total time spent on fetching water and on defecation practice; health outcome (< 5 child diarrhoea incidence as proxy); use of time savings for going to school and effect on enrolment and drop out (particularly girls); effect on productive employment, including use of time savings for productive (self) employment (particularly women); inclusiveness/ distribution of benefits between communities and households.

3. IE coverage

Program components and scale of intervention. As stated above, the program has envisioned delivery of water supply, sanitation and hygiene services for its targeted communities and households and for schools and health facilities. In addition it planned to provide capacity building services to MoWIE, MoH, and MoE and their regional bureaus and Woreda offices who are also involved in the implementation of the program. In this evaluation, the primary focus is on the outcomes/impacts realized at the level of the final target beneficiaries namely, communities and households. At this scale, this evaluation will investigate the impact of all resources invested on the different interventions of the program namely, Water Supply, and Sanitation and Hygiene interventions. Effects of support for facilities provided to schools and health centres will also be studied for a selection of such schools and health centres in or close to sampled communities.

Timeframe. As stated earlier, the implementation of the AfDB-supported program started in 2006 and ended in 2014. However most of the water schemes installed by the program have

started to give services to its target users as from beginning of 2008. The first two years were mainly spent on capacity building at all levels and particularly at district (woreda) level to prepare for selection of beneficiary communities, planning and implementation of water supply, sanitation and hygiene interventions. To give reasonable time for impacts to materialize, the proposed evaluation will cover beneficiary communities with water schemes constructed before July 2013. This time-frame was agreed upon by the Government partners and program staff of the MoWIE who have been consulted in July 2014 during the preparatory mission.

Geographic Coverage. The IE will focus on the three regions with the largest number of beneficiaries from the AfDB supported program and this way it will remain within the available budget for the IE. These regions are Amhara, Oromia and Southern Nations Nationalities Peoples (SNNP) regions in which, respectively, 41, 29 and 24 Woredas out of the total 125 program Woredas in the country are currently found. Together the three regions account for about 87% of the estimated total number of persons benefiting from the program interventions in the country.

Usability of Existing Data Sources. Various sources of secondary survey data (e.g. DHS, WMS and HICE) and administrative databases (e.g. HMIS and EMIS) exist in Ethiopia. However, their suitability to the purpose of the current study is limited. First, the sampling unit used in these surveys is enumeration area (EA) which does not overlap with the primary unit of observation envisaged by the present study (i.e., beneficiary communities of the ADB supported RWSSP). Second, the data from these surveys do not cover the post 2010 period and hence do not provide adequate data to investigate the impact to date of the AfDB-supported RWSSP in the targeted districts. Third, as these surveys have not been specifically designed for measuring the impact of RWSSP, they do not have a sufficient number of beneficiary communities and households in their samples. Fourth, the administrative databases also have limited applications for the envisaged impact evaluation study. From these databases it is difficult to distinguish whether subjects receiving health services from a particular health facility or students enrolled in a particular school are from AfDB-supported communities or not. However, the existing survey databases and administrative records will be used to better understand the trends and patterns of socio economic development over time in the country in general and the study regions in particular. Moreover, reports of these surveys have helped to identify and define relevant indicators for our current evaluation. These have also been used as important sources of parameter estimates needed to compute the size of the evaluation sample.

4. IE questions

The following questions have been set out to steer the IE:

1. What have been the types of water supply, sanitation and hygiene promotion interventions at the level of final target groups ?
2. Was the way the program at the local government and community level evolved as expected?

3. Have communities, men and women, been capacitated to provide safe water supply and basic sanitation and hygiene promotion services?
4. What has been the effect on the access of communities to and their use of improved water supply by households?
5. What has been the effect on human water consumption?
6. Is the drinking water at source and point of use safe (compared to norms)?
7. What has been the effect on the use of toilets -, private, communal, in schools and health facilities? Is open defecation practiced in the village and if so did the extent of this practice change over recent years?
8. What has been the effect on the distance to the primary water source used?
9. Who fetch water from the improved source and what has been the effect on their time spent on fetching water?
10. What has been the effect of interventions on time spent on going to the toilet?
11. Are containers used for transport of water to the homes clean?
12. Do households apply hygiene principles in their water handling at home?
13. Is the water at source and point of use safe?
14. What has been the effect on hand washing practices at critical times?
15. What have been the time savings per day and how have these time savings been used? Which part of time savings have been used for productive (self) employment?
16. What has been the effect of program interventions on productive (self) employment in WaSH service delivery?
17. What has been the effect of the WASH interventions on health outcome variables of interest (taking diarrhea incidence in children-under-five as proxy)?
18. What has been the effect of the WASH interventions on school enrolment and attendance, particularly for girls?
19. Have impacts been similar across communities and households differing in poverty/wealth status, proximity to road or other key distinguishing variables?
20. Are the results sustainable?
 - a. Have institutional structures and roles been well defined and are these understood and fulfilled?
 - b. Have the relevant institutions the capacity to perform the required functions?
 - c. Are the facilities provided technically sound?
 - d. Are the facilities and services provided economically and financially viable, including maintenance and longer term required replacement of infrastructure?
 - e. Do the institutional arrangements provide for adequate monitoring?
21. What has been the role and proportional contribution of the AfDB, as compared to other contributors (communities, governments, development partners)?

5. IE methodology

For this ex post IE a theory based quasi-experimental approach is envisaged. This chapter outlines the methodology to test the ToC and provide credible impact estimates.

Theory of Change:

The study design starts with elaborating the Theory of Change (ToC). A ToC is a description of how interventions are supposed to deliver results, the causal logic of how and why the program interventions will reach intended outcomes. Based on document review and interviews of key stakeholder informants the ToC was reconstructed (schematically presented in annex 1).

Guided by a community demand-driven approach to services delivery the program was expected to overcome critical water and sanitation problems identified and prioritized by target communities by providing different resources and inputs and carrying out several activities and processes to produce water supply, sanitation and hygiene promotion services that are expected to help achieve the desired development outcomes/impacts. Among other things, in order to achieve the desired outcomes, the program interventions involve a series of activities and processes starting with correct identification of beneficiary communities, to community mobilization and training for water supply management and WASH promotion, to training and facilitating local service providers (in areas such as community development, hygiene education, design TA, hand pump installation, construction and maintenance, latrine marketing and construction), providing access to safe water and use of facilities, to improved sanitation and hygiene practices to health and non-health outcomes. In this regard, the study will inspect and carry out factual analysis of empirical data on what actually happened from intervention communities. This analysis will help understand the type and nature of interventions and processes that took place, and what worked and did not work over time in the treatment communities.

The casual link between program outputs and the various expected outcomes/impact involves several pathways and are based on several assumptions. For instance, by expanding access to and use of improved sources of drinking water supply the program has expected to raise women's participation in productive activities and to reduce girls' dropout from school. In the context of rural Ethiopia, women and children, particularly girls bear the burden of fetching water, often traveling long distance to the water source and queuing for several hours.¹ Therefore, installing water schemes closer to homes is expected to release a significant amount of time from fetching water which in turn is expected to benefit most women and girls in the communities targeted by the program. In this respect, the program expects that at least 30% of the time saved is used for productive employment. Also access to a private latrine may save time from finding a safe place for defecating elsewhere.

In the case of children, time saved from fetching water is expected to increase the probability of children's enrollment to school as well as the likelihood of their retention in school once they are enrolled. Likewise, this effect can arise because children are either totally freed-up from the task of fetching water or they require less time to fetch water when the water source is installed closer

¹ Central Statistical Agency (Addis Ababa, Ethiopia) and ICF International (Calverton, Maryland, USA) (2012). Ethiopia Demographic and Health Survey 2011.

to their homes. The education benefits can also result through the health benefits of the program and/or from improved economic status of women.

The program was expected to carry out a series of activities to sensitize communities in the targeted areas on the benefits adhering to proper sanitation and hygiene practices (e.g. use of latrine all the time, proper hand washing at all critical times such as after using a toilet, before eating/feeding, before preparing food). Indeed, evidences from a range of studies suggest that water-borne diseases are directly transmitted not only by drinking contaminated water but also through other 'faecal-oral' routes via microbial agents that can survive on hands and hard surfaces. According to the program's theory, water supply, sanitation and hygiene activities are expected to be integral components of the program interventions in the targeted communities. It is expected that sanitation and hygiene practices are better in areas where water is easily found.

To achieve its objectives the program has assumed the following: First, Woredas are expected to target communities according to criteria set (such as low coverage, poverty, demand and commitment to contribute) and the target communities are expected to have an active role in selection of a suitable water scheme, construction designs and sites for the project. To promote ownership and sustainability of the new investments by the local communities it has also been expected that communities make resource (in cash and in kind) contributions to the construction, operation and maintenance of the schemes. There is a strong assumption that high commitment to this principle exists at all (i.e. community, Woreda, Region and Federal) levels. Communities are also expected to establish water supply and sanitation committees, with 35% women representation, to manage their water systems and promote the building of private latrines and practice hygiene in water handling and hand washing. The Program also expects local private sector to provide and sustain provision of spare parts and maintenance support and other required WaSH services, such as latrine marketing and construction.

Second, it is assumed that the water schemes built are technically sound and appropriate to the local context in order to generate all year-round services to their users. Given the high rate of scheme non-functionality in the country, this is an important assumption. Non-functioning water scheme is a crucial concern because if a water scheme does not work properly or not provide sufficient amount of water, then its users will be forced to seek water potentially from unimproved sources. This situation undermines the potential health benefit of the new initiative.

To what extent these and other related assumptions had been fulfilled in practice will be examined using data collected from the Woreda WASH Team and sample communities and households. In doing so, this study will investigate if these key assumptions have been met.

Impact evaluation design

The fundamental impact evaluation question of this study is 'what would have happened to the outcome variables of interest in communities supported by AfDB's RWSSP had the program not

been implemented?" To answer this question one needs to estimate outcomes for a treatment community in two different states: with and without the program. However, whereas estimating the former is direct because it is directly observed, the latter poses an estimation challenge since is not observed. This is challenge is commonly known as the 'missing data problem' in the impact evaluation literature.

An ideal setting for measuring impact would be to randomly assign some communities to receive the program and keep the rest as a control. In this case, the mean value of the outcome variable for the non-treatment/control group serves as a proxy for the counterfactual outcome for the treatment group had it not participated in the treatment/program. However, in this case, the AfDB program has not been assigned randomly. Instead, the program followed a demand-driven approach and used a combination of different factors to select beneficiary communities. As such, a direct comparison of mean outcomes between treatment and control communities yields biased impact estimates due to potential problems of self-selection and endogenous program placement. Moreover, comparing outcomes before and after the intervention for the treatment communities alone is also misleading because of the assumption that the treatment communities did experience no other changes over time but the program. In practice, this is not a realistic assumption because many factors (e.g. peoples' knowledge, attitude, income, tastes and preferences) that determine the outcome variables, which are subject to evaluation, might be changing over time. As a result, it is difficult to separate effect of general time trends from true program impact. Thus, the 'before-after' impact evaluation method will not be used in this study. Instead quasi-experimental methods will be applied that are commonly applied in observational studies to measure changes caused by the program. In particular, propensity score matching (PSM) methods, (2) difference-in-differences (DD) approach and (3) DD combined with PSM procedure will be used.

The basic intent of the PSM method is to identify control communities which have similar probability of participation in the AfDB-supported RWSSP as with the treatment communities. Accounting for factors of the selection process and a rich set of other relevant community characteristics will be used to match and this approach removes selection bias due to observables. To enhance the internal validity of the evaluation control communities will be carefully chosen during sampling to be as similar as possible with the treatment communities. Furthermore, the control community sample will be further refined during the analysis stage. In particular, in order to reduce bias, treatment/control communities which do not have a comparable propensity score/predicted probability in the other comparison group will be excluded/trimmed out from the estimation sample. Following this, a match will be constructed for each treatment community using an appropriate matching estimator.² Finally, the difference

²There are different matching estimators (e.g. nearest neighbor matching, kernel matching, radius matching etc.) which can be applied. The estimators will be selected that will perform best in terms of matching quality in the context of our evaluation dataset.

in average values of a given outcome variable between the matched pairs of treatment and control communities will be an estimate of the program impact.

In the DD approach, impacts of the program will be estimated by computing the difference in average values of a given outcome variable between the treatment and control communities before and after the program intervention. The measured difference between the two groups will then be tested to determine whether program impact is statistically significant or not. However, the AfDB program did not collect baseline data against which to measure change in outcome variables of interest for treatment and control communities. As such, the DD method will be applied to outcome indicators for which reconstruction of pre-program data is possible through respondent recall. The double-differencing helps to control for time-invariant unobserved community characteristics. To mitigate this problem, the DD approach will be combined with the PSM method.

The PSM and DD analyses estimates mean impacts of the program. The main assumption in estimating the mean impacts is that the treatment effect is equal among the members of the treatment group. In practice, however, a given program can have, for instance, a positive and significant impact for some communities but has a limited or negative impact for others. The differential impacts of a program on its intended beneficiaries might be driven by differences in beneficiaries' existing characteristics. Hence, in this evaluation we will investigate whether or not the impacts of the program vary by beneficiary characteristics such as gender, location (e.g. agro-ecological zone, distance to road and access to town) and poverty (e.g. wealth status, literacy rate) and length of duration of exposure to program interventions.

Outcome variables

The AfDB-supported RWSSP had been designed and implemented to generate positive impacts on a range of water supply and sanitation, health, education, employment, and welfare outcomes. Drawing on several documents of this program and after consultations with key stakeholders of the Water and Sanitation sector in Ethiopia, the following table shows the key outcome/impact indicators for quantitative impact analysis³

Table 1. Definitions of key outcome/impact variables

Household-level outcome variable	
----------------------------------	--

³ Relevant information to determine the minimum required sample size to evaluate impacts on other outcome variables such as hygiene behavior (e.g. hand washing), water quality, and sustainability of water and sanitation facilities have not been available. Hence, program impact on these outcomes may not be precisely measured through econometric analyses. Hence, impacts on these outcomes will be assessed using descriptive analyses.

	Key reference questions in household survey	Community-level outcome indicators
Access and use of improved sources of water supply and sanitation facility		
Using an improved source of water	Does the household obtain its drinking water supply from an improved water source all the time?	Proportion of households obtaining its drinking water from an improved water source; using it all the time?
Travel distance to a water source	How much is the travel distance to reach the primary water source?	Average travel distance to the primary water source (one way)
Total time spent to fetch water	How much time (including waiting, filling time) does it take for the household to fetch water per trip?	Average total time spent per round trip to fetch water
Per capita daily water consumption	How much volume of water is used by a household for domestic use (i.e. drinking, food preparation, cleaning utensils, and basic personal hygiene)	Average daily per capita water consumption
Using a toilet facility	Do household members use a toilet/latrine all the time?	Proportion of households using a toilet facility; using it all the time
Hygiene practices	<p>Do household members wash their hands at critical times (before eating, after going to the toilet, before food preparation)?</p> <p>Is water transport handled safely during transport and storage and use at home?</p> <p>Is water safe at source and point of use?</p> <p>Are toilets clean?</p>	<p>Proportion of households practicing safe hand washing practices</p> <p>Proportion of households transporting water safely to the homes</p> <p>Proportion of households handling water safely at home</p> <p>Proportion of households with clean toilet (without traces of stain)</p>
Health outcome		

Prevalence of diarrhea in under-five children	Was child sick due to diarrhea in the last two weeks	Proportion of under-five children who was sick due to diarrhea in the last two weeks?
Education outcomes		
School enrollment for children	Is a school-age child currently enrolled to school?	Proportion of primary school-age children who are currently registered to attend school (i.e. total, boys, and girls)
School dropout for children	Did child dropout school last year?	Proportion of primary school-age children who dropped out school last year (i.e. total, boys, and girls)
Employment outcome		
Women's participation in productive (self) employment	Was a woman in the household employed in productive (self) employment?	Proportion of women working in productive (self) employment
Local service providers (self) employment	What was the effect on (self) employment in WaSH local service provision	Type and number of local WASH service providers active and approximate hours/days a week

Outcome on access to improved water source is measured using two distinct indicators, namely physical distance to primary water source⁴ and time spent per trip to fetch water. The hypothesis is that a closer water source can lead to a reduced travel time to fetch the water. However, total time spent per round-trip to fetch water can only lead to time savings if time spent at the water source (i.e. queuing and filling) does not also increase.⁵ Use of a latrine/toilet facility is used to

⁴As with Ethiopia Demographic and Health Survey (2011), improved water source is defined to include water sources piped into dwelling/yard/lot, public tap/standpipe, borehole, protected well, protected spring, rainwater, and bottled water.

⁵ Boone, C., Glick, P. and Sahn, D.E. 2011. Household Water Supply Choice and Time Allocated for Water Collection: Evidence from Madagascar. *Journal of Development Studies*, 47(12), 1826-1850.

measure the impact of the program on sanitation coverage.⁶ Water quality, water handling, hand washing practices and cleanliness of toilets are used as variables to estimate impact on hygiene and sanitation practices. The main indicator used to measure the health impact is prevalence of diarrhea in children under-five.⁷ Proportion of women working in productive (self) employment is used as a proxy for the effect of interventions on productive employment and women's economic status. Finally, education benefits will be captured by children's school enrollment rate and school dropout rate. These indicators will also be examined separately for boys and girls to detect if girls have equally benefited from the intervention.

Estimation of sample size

As stated earlier, the AfDB-supported RWSSP had been implemented to generate desirable changes in a number of outcomes/impact variables. However, the sample size estimation focuses on key outcome variables for which the necessary information are available to compute statistical sample. The sample size estimation has followed a number of steps and has made several important assumptions. First, minimization of probability of type-I and type II errors is decided as per common standards namely, 10% (for type I error) and 20% (for type II error). Type I error is the probability of finding a significant program effect when the true effect is zero. On the other hand, the type II error is the probability of not detecting a program impact when it actually occurs. One minus type II error is the statistical power of the test and it is 80% for this evaluation. This means that the design has the power to detect an impact 80% of the times when one has occurred. Second, in the estimations of the minimum number of samples required to measure impact corresponding to each outcome variable of interest, information from the program documents and secondary data sources particularly Welfare Monitoring Surveys, Demographic and Health Surveys, Annual Education Abstracts of the Ministry of Education was used. Third, the unit of assignment of the AfDB-supported RWSSP is the community but most of the outcomes are primarily measured at the level of individuals or households living in those communities. Thus, the sample size estimation followed a two-step procedure and was implemented with STATA Software. Among others, the computation of how many sample communities and households/individuals per sample community need to be surveyed depends on two main factors: (1) the degree of homogeneity/intra-cluster correlation of households/individuals within the community, (2) the survey cost of the evaluation. For instance, when the intra-cluster correlation is high, adding a new household/individual from another community to the evaluation sample gives a greater statistical power than adding a new

⁶ According to Ethiopia DHS (2011) improved sanitation infrastructure is defined to include private access to flush toilet, ventilated improved pit latrine, traditional pit latrine with a slab, and composting toilet. However, in our case we will also distinguish shared access to these facilities and private/shared access to traditional open pit latrine or one without slab from open-defecation.

⁷ Diarrhea is one of the leading causes of child mortality in Ethiopia (FMoH, 2012) as in other developing countries. Existing literature also indicate that about 90% of the times diarrhea is caused by poor water quality and lack of proper sanitation services (Black, 2003).

household/individual from already selected community. Fourth, we applied a balanced design whereby we allocated equal number evaluation sample units between treatment and control groups to maximize the statistical power. Fifth, we anticipate using a two-tailed test in evaluating the program effects so that we will be able to detect both intended and unintended program effects. Finally, we take into consideration the sample size requirement of the proposed impact evaluation method. In particular, we anticipate a 15% trimming of community samples during the process of matching treatment and control communities in the sample. Therefore, the target estimated sample size will be adjusted by 15% inflation factor to obtain the final evaluation sample.

Applying relevant values and important assumptions (as shown in Annex 2), the minimum required sample size for each outcome variable is estimated to be as indicated in Table 2:

Table 2: Estimation of required minimum sample sizes

Outcome/impact indicator	Unit of observation	Minimum required sample size		
		# communities	# households per community	# children per community
Use of improved water source	Household	12	8	
Travel time to water source (one way)	Household	66	4	
Total time spent on fetching water	Household	46	7	
Use of latrine	Household	10	8	
Prevalence of diarrhea	Children aged less than five years	198	12	7
School enrollment	Children aged 6-14 years	40	3	4
	Girls aged 6-14 years	38	5	3
	Boys aged 6-14 years	46	7	5

School dropout	Children aged 6-14 years	210	6	8
	Girls aged 6-14 years	228	12	8
	Boys aged 6-14 years	226	10	7
Women (self) employment in productive economic activities	Adult females age 18-65 years	208	5	

The minimum sample sizes required for the evaluation vary across different outcome variables. However, the minimum sample size required for 'school dropout for girls' variable satisfies the sample size requirements of all the remaining outcome variables. Therefore, our impact evaluation study requires a sample of 228 communities and 12 households per sample community. This implies that with a minimum sample of 114 communities and 1,368 households per comparison group. With this sample it will be possible to detect, for example, a minimum of 30% difference in girls' school dropout rate between the treatment and control communities, with a 90% confidence level and 80% statistical power.

Sampling Strategy

As stated earlier, this evaluation study will be conducted in Amhara, Oromia and SNNP regions. Within these regions it is proposed to choose 38 Woredas in which the program is implemented. The sample Woredas will be allocated to these regions in proportion to the total number of beneficiary people found within them till our evaluation timeframe. Consequently, Oromia, Amhara and SNNP regions will have 18, 12 and 8 Woredas included in the evaluation sample. In each Woreda there will be an equal number of treatment and control community and household samples. In doing so, there will be the same workload for each survey team, and this will facilitate easy field monitoring and supervision, and also ensures self-weighting sample. On the basis of this allocation, the minimum required sample size is distributed across regions as indicated in Table 2:

Table 2. Minimum required evaluation sample size

Region	Treatment sample	Control sample		Total
s				

	Communitie s	Household s	Communitie s	Household s	Communitie s	Household s
Oromia	54	648	54	648	108	1296
Amhara	36	432	36	432	72	864
SNNPR	24	288	24	288	48	576
Total	114	1368	114	1368	228	2736

The next major task is to draw a random sample of treatment and control communities and households within the selected Woredas. To do this, a two-stage selection process will be followed in which the target number of sample communities is selected first and the target numbers of households per selected community are drawn in the second stage. Sampling of communities and households will be performed separately for the treatment and control groups.

Survey Instruments and Data Collection

This study will use Woreda WaSH Team, facilities (water scheme, school, health centre), community-level and household-level survey instruments for primary data collection. The purpose of the brief questionnaire for the Woreda WaSH team is to collect information on what happened in targeting and supporting communities, amount of support provided proportionally by GoE, AfDB and communities, collaboration agreements and effects on number and type of service providers and issues in service delivery. The main purpose of the community-level survey is to gather data on both present and past characteristics of sample communities. The community survey instrument will collect data on several topics including population density, water supply, sanitation and hygiene, public services and infrastructure, proximity to town, natural resources and land use characteristics, agro-ecology features, and labor market information. The non-improved and improved water sources/points used by the community will be mapped and numbered. The household survey instrument will be used to collect detailed information on households' characteristics, use of improved sources of water supply, perception of drinking water quality, use of sanitation facility, hygiene and sanitation practices, time savings and women's use of time savings. In addition water quality at source and at point of use will be tested, the latter for a sample of the sample of households.

The community- and household-level questionnaires will be administered by trained enumerators who speak the local languages of the study areas. The community questionnaire will be administered to a group of five key-informants composed of a water and sanitation committee member, a village elder, two adult females and one adult male members of the community. The main respondent of the household survey will be the household head and his/her spouse. The

field surveys will be carried out in November and December 2014. The surveys will be carried during the same period in the treatment and control sample communities.

There will also be facility-level (water point, primary schools and health centers) surveys to complement the primary data collected through the community-and household-level surveys. As a starting point at the community level in the treatment communities and for a good understanding of what has happened a structured Water Scheme Survey will elicit information on the history of the AfDB supported water scheme and how the WASH services evolved, community contribution to the water schemes, role of WASH committee and knowledge acquired and applied. Water samples will be taken from all primary water sources reported by the household survey respondents in treatment and control communities. Water sample will also be taken at point-of-use/storage at home of about 35% of households who will be chosen randomly from the whole survey households located in the treatment and control communities. Both water samples will be tested to determine water quality at source and point-of-use. The test will primarily focus on *Escherichia coli* (E. coli), which is associated with human faeces. *Escherichia coli* can cause diarrhea and it is commonly used in health studies in developing countries. The water test data will be linked to the household data using the household ID codes. The school survey will primarily record availability, Program provision and use of water supply and sanitation infrastructure in the school and school enrollment and dropout rates. The health facility survey is intended to study effects of facilities provide, if any, and get an idea of the main water-related health problems and associated data on outpatient visits recorded by the health facilities.

A separate follow up mission of the IDEV evaluation team to the surveys and quantitative impact analysis will be fielded to further investigate factors that explain quantitative findings and further address the evaluation questions on sustainability of development results. This part of the study will comprise bottom up data collection from informants of a selection of well performing and less performing treatment communities, private sector, woredas, Regional Bureaus and concerned Ministries.

6. Implementation arrangements

IDEV has overall responsibility for the IE. As mentioned in the introduction the evaluation team will be composed of the IDEV Task Manager, an IDEV Senior Evaluation Officer (co-Task Manager) and a quantitative IE consultant.

Led by the Task Manager the team is responsible for the design and conduct of the IE in accordance with the requirements as per the Approach Paper, inception report and consultancy terms of reference and contract provisions.

Consultancy requirements: The services required from the IE expert concern the quantitative design, quality assurance of the survey, factual and impact data analysis and reporting on the factual and impact analysis and results. These services entail fieldwork, data collection and analysis and participation in report writing and reviewing. In delivering the report the above IE expert is expected to work closely with the IDEV Task Manager and evaluation officer.

The field work, namely the conduction of surveys, will be supported by a national survey firm. These services are commissioned separately.

Quality assurance: The IDEV evaluation team will oversee the field processes to improve the quality of the primary data. After field work, data will be properly edited, cleaned, and entered to computer. Data will be analyzed and interpreted using appropriate models and statistical estimation approach. The result of the quantitative data analysis will inform the follow up mission of the evaluation team.

Reference group: A reference group will be established which comprises the relevant AfDB Operations and Research Department, Field Office and Government of Ethiopia. The group will comment and make suggestions on the expected deliverables. In addition, an external IE expert has been tasked to peer review the IE design, inception report and IE report.

Communication and dissemination strategy: At the time of the AfDB follow up mission to the survey a workshop involving the Field Office and key GoE stakeholders will be held to inform the stakeholders about the findings and obtain feedback. The IE report will be distributed to a wide range of stakeholders and other interested parties within the AfDB and the country.

Work plan

This IE is planned to be completed in April 2015. The study has several phases and will produce different deliverables. To this end, the study will accomplish a number of sequential activities as shown in Table 3:

Table 3: Work Plan of the Impact Evaluation

Phases, Deliverables and Timeline	Proposed time schedule	Responsibility
Phase 1: Inception		
Preparatory mission	21-31 July 2014	IDEV evaluation team
Preparation and refining of inception report and draft survey instruments	20 Aug- 14 Sept 2014	Expert, IDEV team

Submit inception report including Evaluation Methodology and Survey Instruments to GoE	15 October 2014	IDEV
Phase 2: Preparation for Collection of Evaluation Data		
Training of Field Survey Teams	15-20 Nov 2014	Survey Firm, Expert
Pre-testing of Survey Instruments	20-30 Nov 2014	Survey Firm, Expert
Quality assurance mission	1-10 December 2014	IDEV team
Phase 3: Primary Data Collection and Data Entry		
Community, Household and Facility Surveys	1 - 31 Dec 2014	Survey Firm, Expert
Data Editing and Data Entry to Computer	01-25 Jan 2015	Survey Firm, Expert
Phase 4. Data Analysis and Interpretation of Results		
Data Processing and Analysis, and Interpretation of Evaluation Results	26 Jan-20 Feb 2015	Expert, IDEV team
IDEV follow up mission to surveys and assessment of sustainability	1 – 15 March 2015	IDEV team
Writing Draft Evaluation Report chapter on impact analysis	21 Feb-31Mar 2015	Expert
Phase 5. Impact Evaluation Reporting		
Writing/ completing of Draft Evaluation Report	01-15 Apr 2015	IDEV team
Incorporating Feedbacks into the Draft Report	30 Apr – 10 May 2015	IDEV team
Submit Revised and Final Evaluation Report	15 May 2015	IDEV team

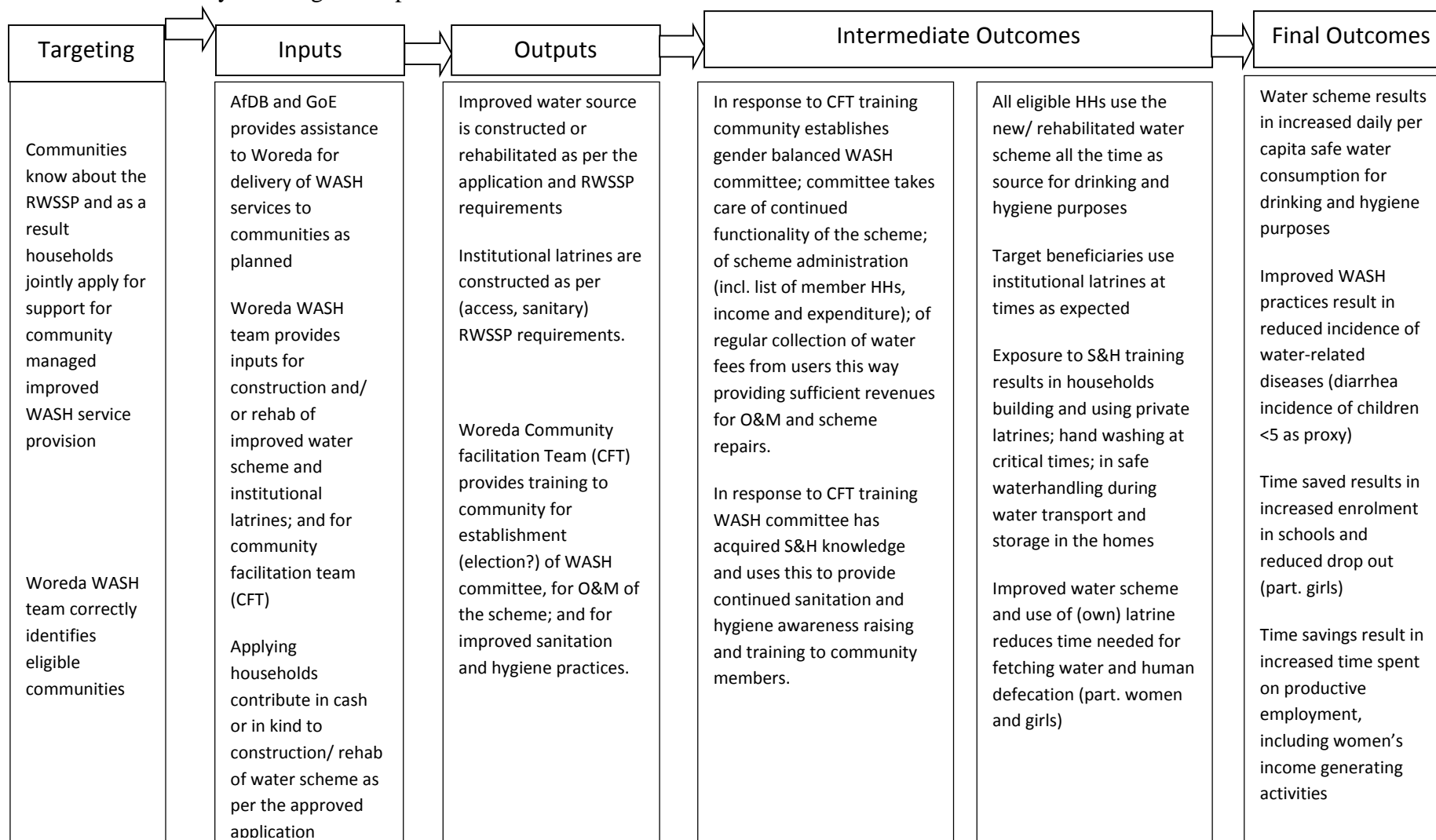
7. Budget

The proposed budget for the study is EU 260,000.00 (equal to UA 220,265.00) exclusive of costs of capacity building and IDEV salary costs. 50% of the budget is funded from the RWSSI Trust Fund.

The breakdown is as follows:

Budget item	Budget in EUROs
IDEV preparatory mission IDEV	11000
Consultancy services IE expert	60000
Survey firm	140000
Survey supervision mission	5000
Follow up field research mission	16000
Dissemination and feedback workshop	10000
Report publication costs	6000
Contingencies	13000
Total	260000

Annex 1: Theory of change Ethiopia RWSSP



Annex 2: Key parameter values used in the estimation of sample sizes

Outcome/impact variables	Base level/rate	Target/minimum effect size	Intra-cluster correlation
Use of improved water source	31.29%	80% target	0.1999
Travel time to water source (one way)	26.28 minutes	Reduction by 67%	0.42074
Total time spent on fetching water	67.64		0.44023
Use of latrine	29.07%	80% target	0.15685
Prevalence of diarrhea	0.18	Reduction by 30%	0.00371
School enrollment: All	0.685	0.9	0.09339
Boys	0.732	0.9	0.07335
Girls	0.63	0.9	0.11727
School dropout: All	0.118	0.0767	0.00894
Boys	0.123	0.07995	0.00251
Girls	0.113	0.07345	0.00761
Women's participation in productive employment	0.5488	Increase by 10%	0.00923

Annex 3: list of reference documents:

1. AfDB Ethiopia, 2005, *Rural Water Supply and Sanitation Program Appraisal Report*
2. AfDB Ethiopia, 2012. *Country Portfolio Performance Review*
3. AfDB Ethiopia, 2011. *Country Strategy Paper*
4. Cairncross, S., Hunt, C., Boisson, S., Bostoen, K., Curtis, V., Fung, I. C. H., & Schmidt, W-P, 2010. *Water, sanitation and hygiene for the prevention of diarrhea*. International Journal of Epidemiology, 39, i193-i205.
5. Federal Democratic Republic of Ethiopia, *Water Supply and Sanitation Program Implementation Manual*, November 2004
6. Federal Democratic Republic of Ethiopia, *Ministry of Health, Community-Led Total Sanitation and Hygiene Implementation Verification Guideline, June 2013*
7. Fewtrell, L., R. Kaufmann, D. Kay, W. Enanoria, L. Haller en J.M. Colford, (2005). *Water, sanitation, and hygiene interventions to reduce diarrhoea in less developed countries: A systematic review and meta-analysis*. Lancet infectious diseases, Vol. 5, pp.42-52.
8. Howard, G., Bartram, J., 2003, *Domestic Water Quantity, Service Level and Health*, WHO/SDE/WSH/03.02.
9. Movik, S., Mehta, L., *The Dynamics and Sustainability of Community-led Total Sanitation: Mapping Challenges and Pathways*, STEPS Working Paper 37, Brighton.
10. NONIE, 2009, *Impact Evaluations and Development, NONIE Guidance and Development on Impact Evaluation*, NONIE - The Network of Networks on Impact Evaluation.
11. OPEV, 2013, *Independent Evaluation Strategy 2013-2017*, Tunis.
12. Prüss, A., Kay, D., Fewtrell, L., Bartram, J. 2002, *Estimating the burden of disease from Water, Sanitation and Hygiene at a Global Level*, Environmental Perspectives, volume 110, no. 5, May 2002.
13. Prüss-Üstün, A., R. Bos, F. Gore en J. Bartram, 2008. *Safer Water, Better Health: Costs, Benefits and Sustainability of Interventions to Protect and Promote Health*. Geneva: WHO.
14. Waddington, H. en B. Snilstveit, 2009. Effectiveness and sustainability of water, sanitation, and hygiene interventions in combating diarrhoea. *Journal of Development Effectiveness* 1(3), pp. 295–335.
15. Waddington, H., B. Snilstveit, H. White en L. Fewtrell, 2009. *Water and sanitation interventions to combat childhood diarrhoea in developing countries*. Study Protocol. Synthetic Review 1 SR 001. International Initiative for Impact Evaluation, New Delhi, March, 2009.
16. WHO, 2002. *The world health report 2002: reducing the risks, promoting healthy life*. Geneva: World Health Organization.

17. WHO/UN-WATER, 2010. *GLAAS 2010. UN-Water Global Annual Assessment of Sanitation and Drinking water supply. Targeting resources for better results*. WHO.
18. WHO/UNICEF, 2010. Joint Monitoring Programme, *Progress on sanitation and drinking water*. Update.
19. World Health Organization (2006). *Guidelines for drinking-water quality; first addendum to third edition; volume 1 recommendations*. Geneva: World Health Organization.