

Tanzania - Transmission and Distribution

Report generated on: March 8, 2018

Visit our data catalog at: <https://data.mcc.gov/evaluations/index.php>

Overview

Identification

COUNTRY

Tanzania

EVALUATION TITLE

Transmission and Distribution

EVALUATION TYPE

Independent Impact Evaluation

ID NUMBER

DDI-MCC-TZA-MPR-TD-2016-v01

Version

VERSION DESCRIPTION

Anonymized dataset for public distribution

Overview

ABSTRACT

The impact evaluation of the transmission and distribution (T&D) activity uses a difference-in-differences approach comparing changes in outcomes over time between the T&D intervention group (villages selected to receive the T&D activity) and a matched comparison group (villages that did not receive the T&D activity). In addition, within the T&D intervention group, we randomly assigned some communities to receive the financing scheme (FS) initiative. To assess impacts of the FS initiative, we compared outcomes of households in the randomly selected treatment communities with those of households in the control communities, adjusting for baseline characteristics. The key research questions we addressed were: (1) What are the impacts of being in communities selected to receive lines through the T&D activity? (2) What are the impacts of being in communities selected to receive low-cost connections through the FS initiative? (3) What are the impacts of being connected to the national grid? (4) Do the impacts vary by gender, age, income, or urbanicity?

Both the T&D analysis and the FS analysis utilize data from a baseline survey (conducted in 2011) and from a follow-up survey (conducted in 2015). The follow-up data are provided in this package. Key findings from the final evaluation include the following:

- The T&D activity led to a large number of new connections, but it was less than a third of the 35,000 connections assumed at the outset.
- The FS initiative also increased connection rates, but even if all T&D communities received low-cost connections, the number of connections originally assumed would still have not been achieved.
- The T&D activity, FS initiative, and connection to the grid increased consumption of grid electricity, ownership of electric tools, time spent watching television, and perceived household safety.
- The T&D activity increased the percentage of households doing an income-generating activity (IGA) that used grid electricity.
- The FS initiative reduced poverty as measured by per capita consumption under \$1 or \$2 per day.
- Being connected increased children's hours of studying at night, but it increased TV watching much more; being connected also increased income and reduced poverty.

More details can be found in the final evaluation report: Chaplin, Duncan, Arif Mamun, Ali Protik, John Schurrer, Divya Vohra, Kristine Bos, Hannah Burak, Laura Meyer, Anca Dumitrescu, Christopher Ksoll, and Thomas Cook. "Grid Electricity Expansion in Tanzania by MCC: Findings from a Rigorous Impact Evaluation." Report submitted to the Millennium Challenge Corporation, Washington, DC: Mathematica Policy Research, February 2017.

EVALUATION METHODOLOGY

Difference-in-Difference

UNITS OF ANALYSIS

For most outcomes, the primary unit of analysis is households. For some community-level outcomes, the unit of analysis is communities.

KIND OF DATA

Sample survey data [ssd]

TOPICS

Topic	Vocabulary	URI
Tanzania Electricity		
Energy		
Electricity		
Africa		

KEYWORDS

Energy, Electricity, Tanzania, Random assignment, Difference-in-difference, Propensity score matching

Coverage

GEOGRAPHIC COVERAGE

The program was implemented in 7 regions across the country, based on their 2011 boundaries: Dodoma, Iringa, Mbeya, Morogoro, Mwanza, Tanga, and Kigoma. The first six regions were included in the baseline survey; the Kigoma region was added to the program later (in the T&D intervention group). The follow-up survey contains data from all 7 regions.

UNIVERSE

For the household survey, the study population was households living in selected communities throughout Tanzania. For the community survey, the study population was selected communities.

Producers and Sponsors

PRIMARY INVESTIGATOR(S)

Name	Affiliation
Mathematica Policy Research	

FUNDING

Name	Abbreviation	Role
Millennium Challenge Corporation	MCC	

Metadata Production

METADATA PRODUCED BY

Name	Abbreviation	Affiliation	Role
Mathematica Policy Research	MPR		Independent Evaluator

DDI DOCUMENT VERSION

Version 1.0 (2016-08-01).

DDI DOCUMENT ID

DDI-MCC-TZA-MPR-TD-2016-v01

MCC Compact and Program

COMPACT OR THRESHOLD

Tanzania Compact

PROGRAM

The Tanzania compact, which ended in 2013, focused on three project areas: the transport sector, the energy sector, and the water sector. The energy sector project focused on four components: (1) the distribution systems rehabilitation and extension activity, also known as the transmission and distribution (T&D) activity; (2) a customer-connection financing scheme initiative to facilitate lower-cost electricity connections in selected areas (hereinafter, financing scheme initiative or FS initiative); (3) promotion of solar power systems in the Kigoma region of mainland Tanzania (Kigoma solar); (4) installation of a new submarine cable connecting Zanzibar's Unguja Island to the mainland along with rehabilitation of various parts of the Zanzibar grid (the Zanzibar interconnector activity, or cable activity). Together, these components were designed to increase the availability of reliable and high quality electricity to people in mainland Tanzania and Zanzibar. This study focused on the first two components.

MCC SECTOR

Energy (Energy)

PROGRAM LOGIC

Through the Energy Activities, MCC financed the design, construction and supervision activities for transmission and distribution (T&D) investments in seven regions. In addition, MCC financed capacity building and technical assistance activities for the implementing entities, including supply of equipment and training sessions for implementing entity. The outputs associated with these inputs include over 1,300 kilometers of 33/11 KV lines constructed, 1,779 kilometers of LV lines constructed, increase in the grid and primary substation capacity, improved policy-related financial sustainability of the utilities, temporary employment through construction contracts, as well as training sessions for implementing entities. The outcomes expected to be realized from the T&D Project include: an increase in the number of domestic, commercial and industrial customers; improvements in the quality of service delivered as measured by reductions in duration and frequency of power outages; increases in the quantity of electricity sold, and reductions in the consumption of other energy sources, such as kerosene and diesel.

PROGRAM PARTICIPANTS

The survey covered households that were selected to be a part of the study, either in the intervention group (villages that received MCC-funded electrical lines) or the comparison group (villages that did not receive MCC-funded lines). Surveys were conducted with as many original participants from the 2011 baseline survey as possible, as well with a new sample of households that had migrated into selected communities or had newly formed from existing households. Household heads (adult males or females in the home) completed the survey, and answered questions about all household members.

Sampling

Study Population

For the household survey, the study population was households living in selected communities throughout Tanzania. For the community survey, the study population was selected communities.

Sampling Procedure

At follow-up, we conducted the community survey in communities located in the original six regions covered by the baseline household surveys, and a seventh region- Kigoma. The follow-up community survey targeted 178 intervention communities and 182 matched comparison communities from the original six regions covered by the baseline household survey. It also targeted 14 new communities in Kigoma.

For the follow-up household survey, we targeted the same 10,908 households that we sampled at baseline, including those that we sampled at baseline but that did not respond to the baseline survey and those that migrated out of the study communities; we excluded 139 households that we sampled at baseline but that we found to be deceased, merged with other households, or duplicates at the time of the follow-up survey. We used this baseline sample for our main analyses of the T&D component and used the intervention group sample for our main analyses of the impacts of the FS component.

In addition, we sampled households in 14 communities in Kigoma that received T&D lines, 2 of which were randomly selected to receive the FS initiative. To create a household sampling frame, we created lists of all households in the sampled Kigoma communities at the time of the follow-up survey. During the listing process, we identified the households residing in the sampled Kigoma communities before October 2011 (at about the time of baseline survey administration in the other regions) and sampled them for the FS evaluation. We sampled a total of 527 baseline households in the Kigoma communities as part of the follow-up household survey.

We also collected data from a sample of "new households" - in-migrant and newly formed households that are new in the study communities since the time of administration of the baseline survey, with communities defined based on their 2011 boundaries. We sampled the new households based on a list of households in each community that we created at the time of the follow-up community survey. We oversampled new households in communities with smaller numbers of new households. We tried to sample, when possible, at least three in-migrant and three newly formed households per community and sample all in-migrant and newly formed households in communities with fewer than three such households. For communities with three or more new households, we also stratified by connection status within the community so that we sampled at least one connected and one nonconnected household when possible and proportionally allocated the sample across these communities and strata unless there was no household in a stratum in that community. We proportionately allocated the sample between the two groups-in-migrant and newly formed households. Out of the total of 2,350 new households that we sampled, we arrived at 1,528 in-migrant and 882 newly formed households from 339 communities, reflecting the share of such households in the study communities. This sample was designed to be representative of the population of in-migrant and newly formed households in the study regions once appropriate weights are applied. However, during the follow-up household survey a large fraction of the households identified as in-migrants or newly formed in the listing told us that they were households in the community at baseline so we ended up not using weights in our analyses.

For more details on sampling and data collection, see the final evaluation report by Chaplin et al. (2017) cited in the abstract above.

Deviations from Sample Design

There were no deviations from the sample design.

Response Rate

Community survey responses were obtained from all 374 communities (the 360 communities from the original 6 regions and the 14 communities in Kigoma).

The response rate for the baseline household survey sample (meaning all targeted baseline households from the 6 original regions) is 81.6% (80.1% in the intervention group and 83.1% in the comparison group). The response rate for the Kigoma sample (meaning households from the 14 Kigoma communities that had been present in the community since 2011, all of

which were in the intervention group) was 84.3%.

The response rate for the new household sample (across all 7 regions) was 74% (73.2% in the intervention group and 75.1% in the comparison group).

Weighting

We constructed four weights to be used in our analyses:

1. FS weight: this was used for the main analyses of the impacts of the FS initiative. This weight covers T&D intervention households that responded to the follow-up survey. It adjusts for sampling (subsidy eligibility) and nonresponse (by region and total migration in the community).
2. T&D weight: this was used for main analyses of the impacts of the T&D activity. It covers all households in the baseline listing that responded to the follow-up survey, but does not cover Kigoma or new households. This is the same as the FS weight for the intervention group. For comparison group households, we used kernel matching to assign weights based on the estimated probability of membership in the intervention group.
3. Exploratory weight: this was used for exploratory analyses of impacts of being connected to the grid. It covers the same households as the T&D weight. The weight is equal to 1 for connected households. For non-connected households, we use kernel matching based on the estimated probability of being connected.
4. Community weight: this was used for analyses of community survey data. It covers the study communities in the original 6 regions (excluding Kigoma). The weight is the size of the community (village or mtaa) as reported by the community leader at baseline.

Questionnaires

Overview

A community questionnaire was administered to leaders in each study community. The questionnaire collected information about the size and composition of the community, availability of key services and infrastructure such as roads and water, access to electricity, availability of civil services such as schools and banks, whether any development projects had been implemented, availability of health services, business activity, and energy/fuel prices.

In addition to the community questionnaire, a household questionnaire was administered to each household in our sample. This questionnaire collected various information on household members such as sex, age, relationship, education, employment, and health. It also collected information on the household's use of electrical and non-electrical energy devices and appliances, income generating activities, consumption and expenditures, use of telephones, time use, assets, income, energy use, and electricity use. In-migrants, newly formed households, and out-migrants were asked additional questions about why they chose to move or form a new household, as well as questions about their income and electricity use prior to moving or forming their household.

At the same time as the community survey, we conducted a household listing to create a sampling frame for new households, and collected GPS data on the location of all electrical poles in survey communities. These datasets are not provided as part of the public use data package but are included in the restricted use package.

The questionnaires were developed in English, and were translated into Swahili. After an initial review the questionnaires were translated back into English by a Tanzania-based researcher. The back translation from the Swahili version was reviewed and compared to the English original. Differences in translation were reviewed and resolved in collaboration with the original translators. The Swahili questionnaires were both piloted as part of the survey pretest.

All questionnaires and modules are provided as external resources.

Data Collection

Data Collection Dates

Start	End	Cycle
2015-05-25	2015-07-11	Community survey, household listing, pole GPS data
2015-08-31	2016-01-27	Household survey

Data Collection Mode

Computer assisted personal interview.

Data Collection Notes

The piloting of the community survey, household listing, and pole GPS data collection was conducted from May 11, 2015 - May 15, 2015. The piloting of the household survey took place from July 15, 2015 - July 22, 2015.

All 374 communities were included in the community survey (360 original study communities from the original 6 regions, plus 14 communities in Kigoma). However, only 371 unique community surveys were conducted, because 3 communities were found to belong to the same village or mtaa as another community.

The community survey took approximately 1 hour to complete.

Most of the household survey was completed by December 5, 2015. In order to improve the response rate, an additional tracking exercise took place from January 7 - January 27, 2016. The household survey required approximately 1 hour to complete.

All data were collected using EDI's in-house software, Surveybe. Interviews were administered in the form of electronic survey questionnaires on hand-held computers.

Questionnaires

A community questionnaire was administered to leaders in each study community. The questionnaire collected information about the size and composition of the community, availability of key services and infrastructure such as roads and water, access to electricity, availability of civil services such as schools and banks, whether any development projects had been implemented, availability of health services, business activity, and energy/fuel prices.

In addition to the community questionnaire, a household questionnaire was administered to each household in our sample. This questionnaire collected various information no household members such as sex, age, relationship, education, employment, and health. It also collected information on the household's use of electrical and non-electrical energy devices and appliances, income generating activities, consumption and expenditures, use of telephones, time use, assets, income, energy use, and electricity use. In-migrants, newly formed households, and out-migrants were asked additional questions about why they chose to move or form a new household, as well as questions about their income and electricity use prior to moving or forming their household.

At the same time as the community survey, we conducted a household listing to create a sampling frame for new households, and collected GPS data on the location of all electrical poles in survey communities. These datasets are not provided as part of the public use data package but are included in the restricted use package.

The questionnaires were developed in English, and were translated into Swahili. After an initial review the questionnaires were translated back into English by a Tanzania-based researcher. The back translation from the Swahili version was reviewed and compared to the English original. Differences in translation were reviewed and resolved in collaboration with the original translators. The Swahili questionnaires were both piloted as part of the survey pretest.

All questionnaires and modules are provided as external resources.

Data Collectors

Name	Abbreviation	Affiliation
------	--------------	-------------

Name	Abbreviation	Affiliation
Economic Development Initiatives Limited	EDI	

Supervision

The community survey, listing, and GPS exercises were led by a team leader and project coordinator with the assistance of a data processing officer and field supervisors. Data collection was conducted by 30 interviewers working in 5 teams of 6, led by 1 field supervisor. On average, a sub-team of 3 completed all 3 exercises within a community in one working day.

The household survey was also led by a team leader and project coordinator with the assistance of a data processing officer and field supervisors. Data collection for the household survey was conducted by 4 teams of 10 interviewers, with each team led by a supervisor. An additional, independent team made up of 8 interviewers were responsible for tracking and interviewing out-migrants for the household survey, and in tracking households that had been classified as unknown or not located in order to help improve the response rate.

To ensure high-quality data collection, supervisors performed direct observations of interviews. This involved supervisors accompanying one of their team members to a community or household and observing the entire interview. The aim was to ensure that the interviewer was conducting themselves in the correct manner, that all questions were being asked correctly and were being explained properly to the respondent(s), and that there was a sufficient amount of probing of respondents. Relevant feedback was provided to the interviewers following the observation exercise.

In addition, the data processing team and coordinator picked a number of communities and households to revisit as a 'spot check' in order to verify the information that was collected during the interviews. A very short interview was conducted, with a subset of important questions included from the actual questionnaires, which aimed to verify a number of important pieces of information in the survey.

Data Processing

Data Editing

EDI performed the first steps of data cleaning. The electronic questionnaires used for the data collection for this survey contained automated routing patterns and a large number of customized consistency checks that provided feedback to the interviewers while collecting data. These primary consistency checks flagged errors inside the questionnaire form and were used by interviewers and supervisors to cross-check collected data in the field. In addition, after the data had been cross-checked in the field by a second interviewer and the supervisor, and transmitted to the EDI headquarters, the data were checked on an ongoing basis by the data processing team using a range of secondary checks in Stata. An additional deeper check was also processed by the Team Leader in order to ensure that all inconsistencies and errors were detected. Whenever the data processing team found errors or inconsistencies, the team contacted the relevant field worker or respondent via telephone to clarify or to identify the correct response, and updated the corresponding interview file. The team furthermore cleaned, streamlined and translated 'other specify' values, open answers and comments, wherever it applied.

EDI submitted cleaned Stata datasets to Mathematica. Mathematica conducted additional cleaning on these datasets. Mathematica conducted additional cleaning of these files in Stata, which included checking the validity of variable values and ranges; verifying skip patterns; cleaning and back-coding common "other-specify" responses; creating binaries of categorical variables; and recoding skips, missing data, and other non-response values to standardized lettered indicators.

Other Processing

All data were collected using EDI's in-house software, Surveybe. Interviews were administered in the form of electronic survey questionnaires on hand-held computers. Automated routing and a large number of built-in consistency checks allowed the identification of errors and missing fields during the interview, while interviewers still had the opportunity to correct the response with the respondent at the source. Electronic interview files were transmitted to EDI's centralized data processing team on a daily basis. The electronic interview files are encrypted, thus interview information could not be accessed by third parties at any point during this project. The data processing team ran additional cross-checks over the data and provided instant feedback to the field teams on an ongoing basis.

Data Appraisal

Estimates of Sampling Error

We did not calculate estimates of sampling error.