
MEMORANDUM**TO:** Jennifer Heintz and Shreena Patel**FROM:** Duncan Chaplin and Arif Mamun**DATE:** 11/23/2015
MCC-TZ 061**SUBJECT:** Quality of data on connections to the national electric grid for evaluation of Millennium Challenge Corporation's energy project in mainland Tanzania

In an effort to promote economic growth and reduce poverty, the Millennium Challenge Corporation (MCC) funded an energy sector project in Tanzania from 2008 to 2013. The Millennium Challenge Account Tanzania (MCA-T) implemented the project. MCC hired Mathematica Policy Research to evaluate impacts of the energy sector project on several outcomes, including connections to the national grid. Mathematica obtained data on connections by community from the Tanzania Electricity Supply Company (TANESCO). In this memorandum, we describe problems with those data and the implications for future work. Comments from MCC and MCA-T staff on earlier drafts of this memorandum and our responses are shown in Appendix A.

A. BACKGROUND

Mathematica planned to use the data on connections by community for estimating impacts of two components of the energy project—the distribution systems rehabilitation and extension activity (also known as the transmission and distribution [T&D] activity) and a customer-connection financing scheme (FS) initiative to facilitate low-cost electricity connections in selected areas.¹ The T&D activity involved rehabilitation of existing electricity transmission and distribution networks as well as construction of new lines in seven regions in mainland Tanzania.² The approximately \$128 million invested in the T&D activity represented more than three-fifths of MCC's total investment in the Tanzanian energy sector project. The FS initiative

¹ Even though transmission lines usually refer to electricity lines of 66 kilovolts or higher capacity, and all electricity lines built under the Tanzania energy project were 33/11 kilovolts or lower capacity, the lines built under the project have been referred to as transmission and distribution lines.

² The seven regions are Dodoma, Iringa, Kigoma, Mbeya, Morogoro, Mwanza, and Tanga. These are the regions based on the 2011 boundaries, many of which have since changed. Kigoma is excluded from the analysis because it was not considered part of the T&D activity at the time of the baseline survey. Consequently, we lack baseline data for that region. We will include Kigoma in some parts of the final report when estimating impacts of the FS initiative.

MEMO TO: Jennifer Heintz and Shreena Patel
FROM: Duncan Chaplin and Arif Mamun
DATE: 11/23/2015
PAGE: 2

was a separate component but closely related to the T&D activity. It was designed to address the concern that normal connection fees present a significant barrier to electricity access for the majority of Tanzanians, particularly for the peri-urban and rural population where the T&D lines were built. Through the FS initiative, approximately 6,000 low-cost connections were made available to residents of a randomly selected set of communities. The FS initiative reduced the connection fee from TZS 320,960 (US \$148) in urban areas and from TZS 177,000 (\$82) in rural areas to TZS 30,000 (US\$14) in both urban and rural areas.³ Implementation of the FS occurred from February 2013 to June 2014, with 1,814 connections (about 31 percent of the available low-cost connections) made under the FS.

For the most part we will be using a subset of the data collected for the evaluation of the T&D activity to evaluate the FS initiative. This is possible because the FS initiative took place in a subset of the communities targeted for the T&D activity. However, we will be excluding communities in the Kigoma region from the T&D evaluation because we did not obtain baseline data on comparison communities without new lines in the Kigoma region. This happened because that region was not part of the T&D activity at the time that we obtained the baseline data. We will include communities in the Kigoma region in the evaluation of the FS initiative because there are communities with and without FS in that region.

The evaluation design for the T&D activity uses a matched comparison group approach that accounts for pre-intervention characteristics (Chaplin et al. 2011).⁴ More precisely, we plan to compare outcomes from 178 intervention communities (that is, communities targeted to receive the MCC-funded line extensions) with outcomes from 182 comparison communities (that is, communities not targeted to receive the line extensions); we matched the comparison communities according to their characteristics before construction of the MCC-funded lines. For the final impact analysis, we will measure the outcomes by using data from follow-up community survey administered between May and July 2015, and household survey administered between September and December 2015. For both analyses, Mathematica is planning to control for pre-intervention characteristics reflected in information collected through the baseline community and household surveys conducted in the intervention and comparison communities in 2011.

Mathematica also planned to conduct an interim impact analysis with the number of connections to the electric grid as the key outcome based on community-level data obtained from TANESCO. We designed the interim impact analyses to provide evidence on short-term impacts of MCC's investments, whereas we expected the subsequent analyses using follow-up survey data to provide evidence on longer-term impact estimates. For both the T&D activity and FS

³ We applied an exchange rate of TZS 2,170 per U.S. dollar (as of June 5, 2015).

⁴ Chaplin, Duncan, Arif Mamun, Thomas Fraker, Kathy Buek, Minki Chatterji, and Denzel Hankinson, "Evaluation of Tanzania Energy Sector Project: Updated Design Report." Report submitted to the Millennium Challenge Corporation. Washington, DC: Mathematica Policy Research, March 2011.

MEMO TO: Jennifer Heintz and Shreena Patel
FROM: Duncan Chaplin and Arif Mamun
DATE: 11/23/2015
PAGE: 3

initiative, Mathematica planned to estimate impacts on the number of connections to the electric grid over the short term—from 6 to 20 months after construction of the T&D lines and from one to 10 months after completion of the FS initiative. This planned analysis would have enabled us to produce impact estimates long before the results from the follow-up surveys would be available (in 2016) as well as to assess how impacts on connections per community change over time. Even though the interim evaluation could have been an important first step, it could not have been used to determine whether the MCC investments in Tanzania's electricity sector were cost-effective. To determine cost-effectiveness, we need longer-term follow-up data to see how households with better access to electricity differ from those lacking such access in terms of health, education, and labor market outcomes. The follow-up surveys conducted in 2015 will obtain such information, and a final report—expected to be available in 2016—will summarize the final estimated impacts of the energy project components.

B. DATA

Mathematica requested data on connections by community from TANESCO for the study communities to estimate the interim impacts of the T&D activity and FS initiative on connections. We obtained community-level administrative data from TANESCO on the number of connections to the national electric grid for each of the 178 intervention communities and each of the 182 comparison communities described above.⁵ The data contained two measures of number of connections for each community—one with the number of connections to MCC-funded lines and another with the number of connections to other (non-MCC-funded) lines. We received connections data at five time points—January, March, June, and December 2013 and June 2014. At the first time point (January 2013), the MCC-funded lines may have been energized in only one of the seven regions covered by the T&D activity (Dodoma). Thus we expected low connection rates to MCC lines at this time point for all regions. The newly built lines in the other regions were energized over time and completed during the period covered by the data. Hence, we expected to see connection rates rise over the time points covered by these data.

To assess the accuracy of the community-level connections data, we relied on data from several other sources. We describe these data sources below and, in the next section of this memorandum, discuss the findings from the quality checks.

- **Baseline household list.** NRECA, a data collection firm contracted by MCA-T, compiled a list of all households in each of the study communities in April and May 2011; the list provided the sample frame for the household survey. The household list includes

⁵ We also obtained data on another 21 intervention communities including 4 that were originally targeted for the T&D lines but were later cancelled, 3 FS communities which were replacements for 3 randomly assigned FS communities, and the 14 in the Kigoma region. In addition, we obtained data on 7 comparison communities that had been selected for the original comparison group but had to be replaced. These additional communities were going to be used in various supplemental analyses.

MEMO TO: Jennifer Heintz and Shreena Patel
FROM: Duncan Chaplin and Arif Mamun
DATE: 11/23/2015
PAGE: 4

information on which households were already connected to the electric grid at baseline or were near enough to the existing electric lines to be able to connect without an additional pole (about 30 meters from the existing lines). We used the data to create two variables. One variable is the total number of households in each community. The other variable is the number of households that were connected to an existing line before implementation of the MCC energy project. Due to natural growth over time in number of connections, we expected to see fewer connections to existing (non-MCC) lines in the data collected in 2011 than in the TANESCO community-level data from 2014.

- **Indicator Tracking Table (ITT) data.** MCA-T collected ITT data on several outcomes at the national level, including total electricity connections to MCC-funded lines in the entire energy project area at the end of second-quarter 2014. We compare the total number of connections from the ITT data to the estimated total based on the community-level TANESCO data for June 2014.
- **Physical house-to-house checks.** Staff from MCA-T and TANESCO and a Mathematica consultant completed house-to-house counts of electricity meters in 55 communities between October 2014 and April 2015. The communities accounted for over two-thirds of the total connections reported in the TANESCO community-level data. We aggregated the data to the community level for comparison with the community-level data from TANESCO.

C. QUALITY OF COMMUNITY-LEVEL CONNECTIONS DATA

Although we estimated interim impacts based on the TANESCO community-level connections data, we subsequently determined that the data were not sufficiently reliable to permit us to publish the findings. Our initial checks of the data suggested that the data were internally consistent and aligned with expected connections numbers found in our baseline data. Hence, we carried out an interim impact analysis based on those data and wrote a draft evaluation report to present the findings. However, after writing the draft report we identified significant discrepancies in the data, as our estimate of the total number of connections in the entire T&D activity area (based on the TANESCO community-level connections data) was found to be more than double the total number of connections reported in the ITT data. We then suggested physical counts of connections by MCA-T and MPR staff to investigate the source of these discrepancies. These additional checks confirmed that the original TANESCO data were inaccurate and unreliable. Knowing that the community-level data over-stated the total number of connections to MCC lines by at least 60 percent, we, in consultation with MCC, concluded that it was inadvisable to publish findings from the interim impact analysis. Given this information MCC decided not to release the interim impact report. Following is a description of the key findings from our checks of the quality of the community-level connections data.

1. Initial checks suggested that quality of the community-level data was promising

Internally consistent community-level connections data

MEMO TO: Jennifer Heintz and Shreena Patel
FROM: Duncan Chaplin and Arif Mamun
DATE: 11/23/2015
PAGE: 5

As discussed, we received community-level connections data for five reporting periods for MCC and non-MCC connections. The data on non-MCC connections were highly correlated across contiguous reporting periods, as was expected. In particular, the non-MCC connections increased at a fairly modest rate, from a total of 13,572 connections in January 2013 to a total of 16,390 connections in June 2014 while the correlations across contiguous reporting periods were nearly 1 for all reporting intervals. This is consistent with reports that TANESCO was working to increase connection rates generally throughout Tanzania during this period, and with the fact that the non-MCC lines had been in place for some time. Thus, very few dramatic changes occurred. Instead there was a slow and steady increase in connections to non-MCC lines for many communities.

While the observed growth in connections to non-MCC lines was slow and stable, the data suggested a very different picture for the MCC lines. For MCC-funded lines, the growth in the connection rates over time was much higher and also much less stable than for the non-MCC lines. Connections to MCC lines increased from a total of only 364 in January 2013 to 9,023 in June 2014, a level of growth that is not surprising in view of the recent installation of MCC lines. Since the lines were installed at different times in each community, we also expected to see higher variability over time in growth rates compared to the non-MCC lines. Indeed, we found that the correlations across reporting periods were substantially lower for the number of connections to MCC lines than for connections to non-MCC lines. The correlations across reporting periods ranged from 0.42 (comparing June and December 2013) to 0.87 (comparing March and June 2013) for the MCC lines. The correlation for the final reporting interval (between December 2013 and June 2014) was also fairly high, at 0.74 which is consistent with the fact that all lines should have been completed by December 2013.

Overall, the pattern of growth in the number of connections over time for MCC and non-MCC lines were consistent with our expectations. The correlations across contiguous reporting periods also seemed reasonable considering the context for each type of line.

Moderate discrepancies with baseline data, but could be due to in-migration

We used the baseline data, collected in 2011, to check for two types of discrepancies with the TANESCO data collected in 2013 and 2014—very large changes in connections to non-MCC lines, and far more total connections (MCC plus non-MCC) than there were households at baseline. For each discrepancy we focused in on those communities where the discrepancies were large both numerically and as a percent of the baseline numbers.

We expected these numbers to differ because of migration and/or changes in existing community members' demand for connections over the two- to three-year period since baseline; so even these discrepancies need not necessarily indicate errors in the TANESCO connections data.

MEMO TO: Jennifer Heintz and Shreena Patel
FROM: Duncan Chaplin and Arif Mamun
DATE: 11/23/2015
PAGE: 6

In 19 percent of the study communities, total connections to non-MCC lines differed by more than 50 connections and by more than 20 percent of the baseline connection counts. In 9 percent of the communities, the total number of connections reported by TANESCO in 2014 exceeded the total number of households from the 2011 baseline list by more than 50 and by more than 20 percent of the baseline counts. Almost all of the latter group of communities was part of the former group; thus, in total, about 19 percent of communities had one or both types of apparent discrepancies. In the remaining 81 percent of the communities, the data were fairly similar and/or plausible. Thus, we viewed the moderate number of discrepancies as reasonable, and considered it as evidence of a need to be prepared for large amounts of in-migration in about one-fifth of the communities included in the evaluation.

Plausible estimated impacts on connections

Using the community-level data, we estimated impacts of the T&D and FS on connections and noted that the estimates seemed reasonable. They were small in January 2013 for both the MCC and non-MCC lines. Subsequently, they grew for the MCC lines and became statistically significant but remained small and generally not statistically significant for the non-MCC lines.

2. Later checks suggested that community-level data are not accurate

Estimate of total connections from community-level data more than double that from ITT data

We found a striking discrepancy when we compared the total number of connections reported in the ITT data (that is, total number of connections to all MCC lines) with an estimate of total connections from the TANESCO community-level data for the baseline household survey communities (that is, total connections to a sub-set of MCC lines).⁶ As of the end of June 2014, the ITT data showed 9,830 total connections for all communities expected to benefit from the T&D activity (except Kigoma) while the TANESCO community-level connections data covering the communities sampled for the T&D baseline household survey, showed approximately 9,023 connections as of the same date. We expected to see less than half as many connections in the study communities compared to all of the T&D communities. Thus, these numbers suggested a major discrepancy.

To estimate the total number of connections to the MCC lines in six of the seven regions (Kigoma being the exception) based on the TANESCO community-level data we had to adjust

⁶ The Kigoma region was excluded from the baseline survey. The ITT data reported 518 connections in Kigoma. The TANESCO community-level data, which was supposed to cover all T&D communities in Kigoma, reported 448 connections in that region.

MEMO TO: Jennifer Heintz and Shreena Patel
FROM: Duncan Chaplin and Arif Mamun
DATE: 11/23/2015
PAGE: 7

for two complications. First, these data covered only one subvillage in each village. Second, they covered only about half of the T&D villages and mitaa for the study.⁷

We dealt with the sub-village issue as follows. Baseline reports from community-leaders suggested that the subvillages surveyed at baseline (those covered by the community-level data described above) would contain about 72 percent of the connections in those villages.⁸ These sub-villages had about 6,338 connections reported by TANESCO. Hence, to estimate total connections for the villages these sub-villages came from we multiplied 6,338 by 1/0.72, which gave us an estimate of 8,803. Adding back the households in communities without multiple subvillages outside of Kigoma (2,685) gives us a total of 11,488 connections in study villages and mitaa.

We dealt with the fact that these data only covered about half of the T&D villages and mitaa as follows. The TANESCO community-level data cover the communities selected for the baseline community survey. We randomly sampled only about 54 percent of all T&D communities for that survey. Hence, we multiply the 11,488 connections by 1/0.54, yielding an estimate of about 21,274 connections. Figure 1 illustrates how the community-level data on connections in the sub-villages and mitaa selected for the survey relate to our estimate of total connections for all T&D villages and mitaa.

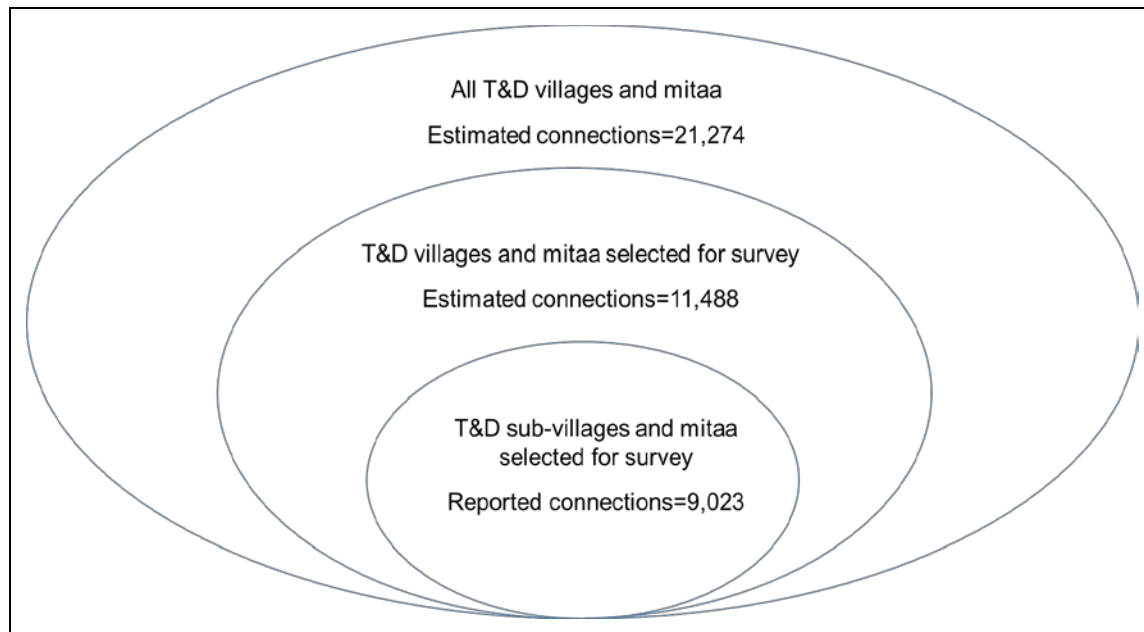
Thus, total connections estimated from the community-level data (21,274) are more than double the number reported in the ITT data (9,023). This discrepancy was not discovered earlier for two reasons. First, until the third quarter of 2014, the ITT data continued to show a reasonably higher number of connections than each iteration of the community-level connections data from the communities included in the evaluation. Second, we did not estimate the total number of connections in *all* T&D communities until we received the last update of the community-level data in June 2014 and had completed our analysis; that's when the stark difference between total connections estimated from the community-level data and the ITT data appeared. Had we developed this estimate of total connections in all T&D communities earlier, it might have been possible to identify the discrepancy a few months sooner, but still long after the initial draft of the interim report was written.

⁷ Mitaa are the smallest administrative units in urban areas in Tanzania. Villages in Tanzania are further subdivided into sub-villages, which are the smallest administrative units in rural areas.

⁸ This is based on a comparison of the reported numbers of households within 30 meters of the new lines in the sub-villages covered by the study compared to the reported numbers for all sub-villages in the villages covered by the study. On average there are 5 sub-villages per village but we selected the sub-villages based in part on the expected number of households within 30 meters of a new line. Consequently we ended up capturing far more than half of these households. The baseline household survey did not cover Kigoma.

MEMO TO: Jennifer Heintz and Shreena Patel
FROM: Duncan Chaplin and Arif Mamun
DATE: 11/23/2015
PAGE: 8

Figure 1: Estimated connections in T&D activity communities based on community-level data



Physical house-to-house checks found far fewer connections than community-level data

After we noticed the discrepancy in total connections from the ITT and community-level data, staff from MCA-T and TANESCO and a Mathematica consultant performed physical house-to-house counts in 55 study communities. They found a total 3,814 connections in those communities compared to the estimated 6,188 connections based on the TANESCO community-level data. Thus, the community-level data yielded counts about 62 percent higher than the physical counts.

3. No clear evidence of biased reporting

We found no evidence of biased reporting in the community-level connections data. While the estimate of total connections based on the community-level data (21,274 connections) is more than double what the ITT data suggest, it is less than two thirds of the target of 35,000 set in MCC's 2008 economic rate of return (ERR) calculations. This means that the ITT data were far from being high enough to show that the target was achieved. In addition, whereas 39 of the 55 communities subjected to house-to-house checks showed overestimates of connections in the community-level data, another 12 communities had too few connections (that is, the community-level data suggested fewer connections than found in the house-to-house counts), and 4 had the

MEMO TO: Jennifer Heintz and Shreena Patel
FROM: Duncan Chaplin and Arif Mamun
DATE: 11/23/2015
PAGE: 9

exact same numbers by both counts.⁹ Finally and perhaps most important, the same TANESCO staff who provided the community-level data also provided the ITT data that initially alerted us to the problem with the community-level data. For all of these reasons we do not feel that these data provide compelling evidence of intentional misreporting.

There are several possible explanations for the patterns observed in these data that do seem more plausible than intentional miss-reporting. First, during MCA-T/MPR verification trips to district offices, TANESCO staff reported that they often made rough estimates for the connections reports. Second, in some cases, it appears that TANESCO staff may have based their counts for the community-level reports on data associated with subprojects or transformers rather than with communities (subvillages or mitaa). Subprojects generally cover several communities and thus using subproject data might easily lead to general overestimation at the community level. Finally, community boundaries and names changed considerably over the years, and TANESCO often used region and district names that differed from those used for general administrative purposes. Consequently, it would be easy for a local TANESCO officer to be unclear regarding which communities were supposed to be included in their reports.

D. IMPLICATIONS

1. Implications for the energy project evaluation

Our analyses suggest that the TANESCO community-level data overstate the number of connections to MCC lines by at least 62 percent and perhaps by more than 100 percent. As discussed, Mathematica used the community-level data to produce a draft interim report with estimated impacts of the T&D activity and FS initiative on community-level connections. Even though the results of such an analysis would have been helpful for policymakers, the inconsistencies we identified in the data suggest that the results might be misleading. Hence, we recommended and MCC agreed not to publish the interim report. Fortunately, we expect to be able to obtain more accurate data on connections by community from the community and household survey data that we are collecting in 2015. We will use those data to estimate impacts on community-level connection rates in a report that is expected to be finalized in 2016.

Our analyses of the community-level data from TANESCO were also supposed to be used to inform the timing of the follow-up community and household surveys. Once we learned that the community-level data from TANESCO were not reliable, we made the decision to work with MCA-T and TANESCO to obtain the additional data based on house-to-house counts to help inform our decisions about the timing of the community survey. This required a great deal of

⁹ We suspect that, in at least some cases, the community-level numbers were too low because of natural increases in connections between June 2014—when the community-level data were collected—and October 2014 to April 2015—when the physical house-to-house checks were performed.

MEMO TO: Jennifer Heintz and Shreena Patel
FROM: Duncan Chaplin and Arif Mamun
DATE: 11/23/2015
PAGE: 10

extra effort and resources. In the end we decided to conduct the follow-up community survey and list all households in the sampled communities during May through July 2015.

Based on our analyses of the TANESCO community-level data we were also planning to implement the follow-up household survey during September through December 2015; given the data problems encountered we ended up using data from the follow-up community survey and the household lists to make a final recommendation on the timing of the follow-up household survey, which was to continue with conducting the follow-up survey in the fall of 2015. In making recommendations for timing of the household survey, we assessed the number of connections to the national grid in the T&D intervention and comparison communities as that was a critical intermediate step for electrification to directly affect changes in household outcomes. Had the TANESCO community-level connection data been reliable, the decision about the timing of the follow-up household survey could have been made sooner and with greater confidence. Since the TANESCO Community-level data were not reliable we were at some risk of additional costs if we ended up having to postpone the household survey to be many months after the community survey. In the end, while the poor quality of the community-level connection data did not derail the evaluation, it did put the evaluation at heightened risk in terms of the timing of the follow-up household survey.

2. Implications for future evaluations

We would like to point out a few lessons learned from our experience in obtaining administrative data on the number of connections by community. First, our results suggest the need for greater care when collecting community-level data on connections. In particular, when asked to provide such data TANESCO should clarify what they can and cannot provide, and then work to make sure that they do provide the information promised with a reasonable degree of accuracy. We understand that TANESCO does not have dedicated monitoring and evaluation staff. Identifying key staff to play this role on future projects of this magnitude may be helpful. Second, it would also be useful to encourage TANESCO staff to propose alternative solutions when asked to collect data that they do not have readily available. Third, TANESCO may want to invest in developing a reliable and comprehensive customer-level database covering all customers connected to the Tanzanian national grid, and to consider building capacity among field level engineers and technicians to collect and record the relevant data accurately. Such data will likely be useful for billing and maintenance, as well as for future evaluation work. Fourth, MCC may want to encourage MCA-T or the evaluator to conduct more physical on-the-ground checks of data to ensure that the data match what is reported. Fifth, when possible, it would be helpful to align the study units with units used by TANESCO upfront and agree to using this shared set of units and their respective names/codes going forward. Finally, in the future, it may be possible to obtain GPS location data on new TANESCO customers at the individual level, thereby alleviating problems associated with changes in community names and boundaries.

MEMO TO: Jennifer Heintz and Shreena Patel
 FROM: Duncan Chaplin and Arif Mamun
 DATE: 11/23/2015
 PAGE: 11

APPENDIX A. STAKEHOLDER COMMENTS AND MATHEMATICA RESPONSES

Reviewer Name/ Institution	Page Number (please reference the number at the bottom of the page)	Comment	MPR response
MCA-T energy sector	General	It is to be pointed out also as lessons learned that Generally TANESCO does not have a Monitoring and Evaluation team, department with M&E staff the data collection is not assigned to specific responsible staff who knows what is needed and for what reason. Their reporting format has a different data need compared to what was needed for this aspect. Dedicated M&E staff were needed from onset (i.e. during Baseline studies).	We have added a discussion of the need for dedicated M&E staff at TANESCO in the implications section of the memo. The memo already contained the following recommendation to address the concern about the reporting format: "...when possible, it would be helpful to align the study units with units used by TANESCO upfront and agree to using this shared set of units and their respective names/codes going forward."
MCA-T energy sector	Foot note 1:	Although changes came later after the Baseline studies. It is worth mentioning that our scope of work now includes new administrative regions; Iringa region has been split to include Njombe region, Mwanza has been split to include Geita region and Dodoma has Manyara region and in all these new regions we have T&D scope of work. The intervention eventually covers 10 regions.	We have augmented footnote 2 to discuss changes in the boundaries of the regions.
MCC Economic Analysis	6	The term "mitaa" is used on this page without being introduced previously, although it is described later in the memo. Please move the description up to coincide with the term's first use.	We have added the following sentences in a new footnote at the end of the first sentence that uses the word "mitaa": "Mitaa are the smallest administrative units in urban areas in Tanzania. Villages in Tanzania are further subdivided into sub-villages, which are the smallest administrative units in rural areas."

MEMO TO: Jennifer Heintz and Shreena Patel
 FROM: Duncan Chaplin and Arif Mamun
 DATE: 11/23/2015
 PAGE: 12

Reviewer Name/ Institution	Page Number (please reference the number at the bottom of the page)	Comment	MPR response
MCC Economic Analysis	General	<p>Would it be possible to include a few charts and graphs that illustrate the issues a little better? If we are going to post this on MCC's website, it may be a little difficult for laypeople to follow all of the numbers a little more easily. Something like the number of estimated connections by each method over time might help people see the discrepancies a little better. Some sort of graphical explanation of how the estimates were reached (multiplying by 1/0.72, etc.) might improve comprehension too.</p>	<p>We have added a new figure to help illustrate how the community-level data on connections relates to our estimate of total connections in all T&D communities. We did not add a figure showing connections over time but have added a related discussion. See our response to the next comment for details.</p>
MCC Economic Analysis	General	<p>Given the data quality issues, the decision not to publish the Interim Connections Report is understandable. However, the memo does not make clear why the discrepancies were not caught before the report was written. Were the errors noticed during the review process for the interim report, or was there something else that triggered the additional scrutiny on the connection numbers? If we want to avoid a situation like this again, we should think about how to detect these issues as early as possible. The memo suggests some prevention measures, but some suggestions for additional data quality checks at each stage in the process would be helpful too.</p>	<p>We have added a discussion of why the data problems were not discovered earlier on p. 7, at the end of the section, "Estimate of total connections from community-level data more than double that from ITT data".</p>

MEMO TO: Jennifer Heintz and Shreena Patel
 FROM: Duncan Chaplin and Arif Mamun
 DATE: 11/23/2015
 PAGE: 13

Reviewer Name/ Institution	Page Number (please reference the number at the bottom of the page)	Comment	MPR response
TANESCO	Implication - for future evaluations, part 2, page 9	We agree with all suggestions and proposals given by the writer. To add on, 1) There is a need for capacity building to TANESCO staff, especially at lower level (Engineers and Technicians) who in most cases involve directly in the field for data collection and data entry to realize the importance of accuracy data in reporting. 2) There is a need of assisting TANESCO to have a very reliable database which could be used to keep all targeted data. Currently TANESCO does not have a centralized database.	We have added text on page 10 suggesting potential benefits of having TANESCO create and maintain a more complete and accurate customer-level database.
Ministry of Finance, Poverty Eradication Department		No comments.	