



MCG Data Quality Review

Final Report

Round IV

February 23, 2011

Prepared for:

Millennium Challenge Georgia
(MCG)

Prepared by:

IMPAQ International, LLC
10420 Little Patuxent Parkway,
Suite 300
Columbia, MD 21044
Telephone: (443) 367-0088
Facsimile: (443) 367-0477

MCG DATA QUALITY REVIEW

FINAL REPORT

TABLE OF CONTENTS

Executive Summary	i
Chapter 1. The Data Quality Review Study	1
1.1 Introduction.....	1
1.2 Purpose of Data Quality Review.....	1
1.3 Data Quality Review Methodology	2
1.4 Data Received from MCG	3
1.5 Organization of the Report.....	4
Chapter 2. Performance Monitoring Indicators	5
2.1 Introduction.....	5
2.2 Methodology for Rating Performance Indicators	6
2.3 Ratings of Performance Indicators.....	11
Chapter 3. Review of the Surveys.....	14
3.1 Introduction.....	14
3.2 Survey Quality Ratings	15
3.3 ADA Beneficiary Survey.....	17
3.3.1 Overview of the ADA Beneficiary Survey.....	17
3.3.2 Brief Description of Findings from ADA Surveys.....	18
3.3.3 General Comments on the ADA Survey.....	24
3.4 Settlements Infrastructure Survey (SIS)	26
3.4.1 Overview of the SIS.....	26
3.4.2 Brief Description of Findings from SIS Round IV Survey (2010).....	26
3.4.3 General Comments on SIS.....	26
3.5 Road Users Survey.....	27
3.5.1 Overview of the Road Survey.....	27
3.5.2 Brief Description of Findings from the Road Survey.....	28
3.5.3 General Comments on the Road Survey.....	28
3.6 Village Infrastructure Census (VIC).....	29

3.6.1	Overview of the VIC.....	29
3.6.2	Brief Description of Findings from VIC.....	30
3.6.3	General Comments on VIC.....	31
3.7	Integrated Household Survey (IHS).....	32
3.7.1	Overview of the IHS	32
3.7.2	Brief Description of Findings from IHS	33
3.7.3	General Comments on IHS	33
Chapter 4. Summary and Conclusions		34
Appendix A: Rating Tables for Program Performance Indicators.....		36
Appendix B: Rating Tables for Surveys		53

LIST OF TABLES

Table 1.1: Survey Data Received from MCG.....	3
Table 2.1: Performance Indicators for Round IV.....	5
Table 2.2: Instrument Validity Sub-Criteria.....	9
Table 2.3: Rating of Average Annual Daily Traffic (AADT) Indicator.....	10
Table 2.4: Summary Ratings for Round IV Performance Indicators.....	12
Table 3.1: Summary Ratings for Round IV Data	16
Table 3.2: Overall Ratings of Surveys, Rounds I-IV	16
Table 3.3: Sample Characteristics, ADA Surveys.....	18
Table 3.4: Gender Distribution, Primary Producer Survey Respondents	18
Table 3.5: Gender Distribution, Value Adders- Value Chain Survey Respondents.....	19
Table 3.6: Gender Distribution, Farm Service Center Survey Respondents	19
Table 3.7: Regional Distribution, Primary Producer Survey Respondents	21
Table 3.8: Regional Distribution, Value Adders- Value Chain Survey Respondents	22
Table 3.9: Regional Distribution, Farm Service Center Survey Respondents.....	23
Table 3.10: Ability to take a loan, Primary Producer Survey Respondents	24
Table 3.11: Ability to take a loan, Value Adders-Value Chain Survey Respondents.....	24
Table 3.12: Ability to take a loan, Farm Service Center Survey Respondents	24
Table 4.1: Summary Ratings for Round IV Data	34
Table 4.2: Overall Ratings of Surveys, Rounds I-IV	35
Table A.1: Rating of Poverty Gap in the Samtshke-Javakheti Region Indicator.....	36
Table A.2: Rating of Poverty Incidence in the Samtshke-Javakheti Region Indicator.....	37
Table A.3: Rating of Jobs Created from Enterprise Development Activities Indicator	38
Table A.4: Rating of Household Net Income Indicator.....	39
Table A.5: Rating of Savings in Vehicle Operating Costs (VOC) Indicator	40
Table A.6: Rating of International Roughness Index (IRI) Indicator	41
Table A.7: Rating of Annual Average Daily Traffic (AADT) Indicator.....	42
Table A.8: Rating of Travel Time Indicator	43
Table A.9: Rating of Increase in Gross Revenue of Portfolio Companies (PCs) Indicator .	44
Table A.10: Rating of Increase in PC Employees Indicator	45
Table A.11: Rating of Increase in Local Supplies to the PCs Indicator	46
Table A.12: Rating of Increase in Wages Paid to the PC Employees Indicator	47
Table A.13: Rating of Increase in Locally Sources Goods and Services Paid Purchased by PCs Indicator.....	48

Table A.14: Rating of Jobs Created Indicator	49
Table A.15: Rating of Household Net Income Indicator.....	50
Table A.16: Rating of Household Firm Income Indicator	51
Table A.17: Rating of Beneficiaries Indicator.....	52
Table B1: ADA Direct Beneficiary Survey	53
Table B2: Settlement Infrastructure Survey (SIS)	54
Table B3: Road Users Survey	55
Table B4: Village Infrastructure Census (VIC).....	56
Table B5: Integrated Household Survey (IHS).....	57

EXECUTIVE SUMMARY

In September 2005, the Millennium Challenge Corporation signed a five-year, \$295.3 million Compact with the Government of Georgia. In November 2008, MCC and the Government of Georgia signed a Compact amendment making up to \$100 million of additional funds available; thus making the total grant value \$395.3 million. The main objective of the grant is to facilitate poverty reduction through economic growth in Georgia. To achieve this goal, the Compact will support programs to eliminate development constraints through infrastructure rehabilitation, roads improvement, and investments in SMEs and agribusinesses.

There are two major projects within the Georgia Program:

- 1) **Regional Infrastructure Rehabilitation (RIR) Project** – The objective of the RIR project is to rehabilitate key regional infrastructure through three activities:
 - **Samtskhe-Javakheti (S-J) Road Rehabilitation** – Rehabilitate and construct more than 220 kilometers of a main road traversing the S-J region.
 - **Main Gas Pipeline Rehabilitation** – Rehabilitate the North-South gas pipeline that fuels electric power generation and provides heat to homes and businesses, and develop the Georgian government energy sector strategy.
 - **Regional Infrastructure Development (RID)** – Fund regional and municipal physical infrastructure to improve potable water supply.

- 2) **Enterprise Development Project** – The objective of this project is to promote regional enterprise development through investments in SMEs and agribusinesses through two activities:
 - **Georgia Regional Development Fund (GRDF)** – Fund a professionally and independently managed investment fund to provide long-term capital and technical assistance to SMEs and to identify legal and policy reforms needed to improve the investment environment.
 - **Agribusiness Development Activity (ADA)** – Provide grants and technical assistance to farmers and agribusinesses that supply agricultural products to the domestic and international market.

Millennium Challenge Georgia (MCG) has engaged various organizations to gather data that are required to report progress and to monitor and evaluate Program activities. These data are used to monitor implementation progress, to measure whether the Program is achieving its intended results, and to make programmatic adjustments as necessary. Furthermore, these data are also being used to conduct rigorous impact evaluations of specific activities.

Because of the importance of high quality data for effective monitoring and evaluating of the Program, MCG contracted IMPAQ International in February 2009 to verify the quality and consistency of the data over time, across different implementing units, and other reporting institutions. IMPAQ's Data Quality Review (DQR) has been conducted in four rounds:

- Round I - all data through the end of PY2;
- Round II - all data through the end of PY3;
- Round III - all data through the end of PY4; and
- Round IV - all data through the quarter 3 of PY5.

During each round, IMPAQ has assessed the quality of the data that has been collected and identified areas where the data quality is appropriate and where the data quality is in need of improvement. The goal of the DQR activities is to ensure that data reported are as valid, reliable, timely, and precise as resources allow. As part of the data quality reviews, staff identified key issues and/or problematic areas regarding data quality and mitigation measures to correct the problems. In addition, staff advised the M&E Unit within MCG on the appropriate timing and approach for conducting spot checks of data quality through field visits.

This report presents the results of the DQR activities of Round IV. Specifically, the report provides an overview of the objectives of the project and the available data (Chapter 1), the program performance indicators for Round IV and the methodology for assessing their quality (Chapter 2), a review of the quality of the Round IV data and a summation of the entire DQR process (Chapter 4).

Overall, the conclusion of Round IV of the DQR is that the M&E Unit within MCG has made substantial progress in improving the accuracy and reliability of the data that is now collected for monitoring and evaluating the Program. In particular, all recommendations that were made by

IMPAQ in earlier rounds have now been implemented and MCG's data collection activities now appear to be of high quality. This willingness to accept recommendations for improving data collection activities has resulted in steadily improving the data accuracy and reliability over time.

CHAPTER 1. THE DATA QUALITY REVIEW STUDY

1.1 Introduction

The Millennium Challenge Georgia (MCG) Fund selected IMPAQ International and its partner, the Policy and Management Group of Georgia (PMCG), to review the quality of program monitoring and evaluation (M&E) data. As stated in the contract between MCG and the IMPAQ Team (IMPAQ and PMCG), the Data Quality Review (DQR) will provide critical information for program management and decision-making and also on progress achieved toward the Program's objectives and goals. The DQR will cover all data collected during the five-year life of the Compact between the Millennium Challenge Corporation (MCC) and MCG.

This report presents an analysis of the data quality of five primary surveys and their respective data sets. The report also presents select performance indicators chosen for review during Round IV of the DQR. In addition to presenting the results of a review of Round IV data, the report also examines the changes in data quality that have been observed by the reviewers over time.

The IMPAQ Team's overall conclusion from Round IV of the DQR is that MCG continues to improve the overall quality of the data that it is collecting, analyzing, and reporting. This conclusion is supported by the higher rating scores for the indicators assessed in this report as well as by the observation that all of the recommendations made by the IMPAQ team have been adopted by the organizations responsible for the data collection. Furthermore, the adoption of our recommendations was generally made thoughtfully, and often the implementing organizations went beyond our recommendations and sought additional approaches to improve data quality. It appears that MCG continues to make significant progress in increasing the data quality and in its reporting efforts to MCC.

1.2 Purpose of Data Quality Review

The objective of the DQR is to ensure that the data collected for Program M&E are of acceptable quality, reliability and consistency. The overall goal is to verify the quality and consistency of data across different Implementing Entities (IE) and other institutions engaged in surveys. The

DQR will also assist in identifying key issues or problematic areas regarding data quality and identifying mitigation measures to correct the problems.

As part of this DQR for Round IV, the IMPAQ Team will assess the following:

- Methodology used to collect the data and the consistency with which the methodology is being followed in ongoing data collection;
- The accuracy, consistency, and reliability of primary and secondary data;
- Data dissemination and confidentiality;
- The validity of the instruments used for data collection (i.e., are the surveys measuring what they intend to measure);
- Survey quality and accuracy;
- The methodology and accuracy of reporting to MCC;
- The completeness of the data (i.e., were the needed data fully collected or are there gaps?);
- The availability of adequate metadata and documentation for surveys that allows correct use/understanding of the data sets;
- The comparability and consistency of data collected within and across data sources;
- The ability to combine/merge/link different data sources for analysis purposes;
- The accuracy of data handling, data cleaning, derived variables, aggregation at different levels, as well as handling outliers and missing data across data sources;
- Adequacy of the data and methodology for making inference/generalizing results at different level (nationwide, regional, district, etc.); i.e., can the data be used to make inferences about the population of interest?

Additionally, the DQR will cover all data reported in the M&E Plan including monitoring data submitted by IEs and any survey data collected by partner organizations financed by MCG.

1.3 Data Quality Review Methodology

IMPAQ conducted the following reviews for Round IV of the DQR and will, in this report, provide feedback to MCG on relevant findings from our review of the following:

- M&E plan
- Surveys

- Data collection process
- Data input process
- Data analysis process
- Reporting system.

While some of the DQR activities (e.g., observational analyses of the data collection processes) have been implemented in Georgia by consultants, much of the DQR has been conducted through desk reviews of project documents and analysis of project data.

1.4 Data Received from MCG

Throughout this project, MCG provided IMPAQ with relevant documents and data files of the survey data collected in each of the four Rounds. Table 1.1 provides a summary of the survey data the IMPAQ team received throughout this project. As shown in Table 1.1, the IMPAQ Team has received Round IV data from all surveys, except for the RID beneficiary survey. In previous reports, we examined the quality of the survey data collected in Rounds 1, 2, and 3. This report focuses on examining the quality of each survey data in Round IV and how it compares to the quality of the survey data in previous rounds.

Table 1.1: Survey Data Received from MCG

Survey Data	Round I	Round II	Round III	Round IV
Integrated Household Survey (IHS)	✓	✓	✓	✓
Village Infrastructure Census (VIC)	✓	✓		✓
Agribusiness Development Activity (ADA) Beneficiary Survey	✓	✓	✓	✓
Settlement Infrastructure Survey (SIS)	✓			✓
Regional Infrastructure Development (RID) Beneficiary Survey			✓	
Road Users Survey			✓	✓

1.5 Organization of the Report

We organized the remainder of this report as follows. Chapter 2 provides an overview of the 17 Round IV program performance indicators and the process used to assess their quality. Chapter 2 also presents the ratings of each performance indicator based on seven criteria: validity, reliability, timeliness, precision, objectivity, adequacy and practicality. These criteria are in turn used to assess the quality of the indicators. Chapter 3 examines the quality of the Round IV data in detail, including: 1) an assessment of the quality of the information reported in the data, 2) comparison of data quality relative to the quality of previous rounds of data, and 3) overall assessment of data quality based on the aforementioned criteria. Chapter 4 summarizes the findings of the report.

CHAPTER 2. PERFORMANCE MONITORING INDICATORS

2.1 Introduction

To measure program performance, MCG needs appropriate performance monitoring indicators and appropriate data sources for measuring the indicators. Throughout this project, the IMPAQ Team assessed the quality of the performance monitoring indicators that have been developed as well as the quality of the data collected for measuring these indicators. The 17 indicators are presented in Table 2.1 (2 are Goal Level indicators; 2 are Objective Level indicators; and 13 are Outcome Level indicators). The remainder of this chapter provides a description of the methodology used to assess the quality of these indicators and the results of the assessment.

Table 2.1: Performance Indicators for Round IV

Goal Indicators	Source
Poverty gap in the Samtshke-Javakheti Region	DS
Poverty Incidence in the Samtshke-Javakheti Region	DS
Objective Indicators	
Jobs created from Enterprise Development Activities	MCG
Household Net Income	MCG
Outcome Indicators	
S-J Road Rehabilitation	
Savings in Vehicle Operating Costs (VOC)	GORBI
International Roughness Index (IRI)	RDMED
Annual Average Daily Traffic (AADT)	GORBI
Travel Time	GORBI
Georgia Regional Development Fund	
Increase in gross revenues of the Portfolio Companies (PCs)	SEAF
Increase in PC Employees	SEAF
Increase in local Supplies to the PCs	SEAF
Increase in wages paid to the PC employees	SEAF
Increase in locally sourced goods and services purchased by PCs	SEAF
Agribusiness Development Activity	
Jobs Created	CNFA
Household Net Income	CNFA
Firm Income	CNFA
Beneficiaries	CNFA

2.2 Methodology for Rating Performance Indicators

To assess each indicator, we defined a set of criteria based on international sources that define assessment criteria for program performance indicators. The primary source for these criteria is the OECD/DAC Glossary of Key Terms in Evaluation and Results Based Management,¹ a document developed by the Development Assistance Committee (DAC) Working Party on Aid Evaluation of the OECD. This Committee consisted of high-level officials and academics in the international development community, including bilateral and multilateral development evaluation experts from the World Bank, UNDP, IDB, and the community of international donors from 30 member countries. Additional sources included: 1) the Evaluation Thesaurus (Fourth Edition) by Michael Scriven; 2) Evaluation (Second Edition) by Carol Weiss; and 3) USAID's Center for Development Information and Evaluation *Performance Monitoring and Evaluation Tips*.²

Based on the recommendations provided by these sources, we have defined 7 criteria for assessing the quality of the performance indicators:

- **Validity** – This criterion examines the extent to which to which the indicator captures the concept of interest. Specifically, an indicator is considered valid if it satisfies the following conditions:
 - The relationship between the indicator and the program exists.
 - Indicator is defined well and reported in an appropriate format in the data.
 - Indicator values accurately represent reality.
 - Indicator has a verifiable source.
 - Indicator presents all necessary data.
- **Reliability** – This criterion examines the consistency and dependability of the indicator with regards to the quality of the instruments and process used to collect and report the information. An indicator is considered reliable if it satisfies the following conditions:
 - Data is collected in a consistent way over time, across project activities, and regions.

¹ <http://www.oecd.org/dataoecd/29/21/2754804.pdf>

² http://pdf.dec.org/pdf_docs/pnaca927.pdf

- There is a Quality Control mechanism ensuring that the data collection and reporting process is designed to guard against bias, the process is periodically reviewed, and that random checks are performed at each stage of the process.
- The process for collecting and documenting the information is documented and problems are reported in a transparent way.
- **Timeliness** – This criterion examines if the indicator was collected and reported regularly based on the relevance of the indicator. The timeliness of an indicator is assessed based on the following conditions:
 - The data is collected on a frequent basis, enabling its use to regularly inform program management decisions.
 - The data is sufficiently up to date to facilitate decision-making.
 - The date of data collection is clearly identified.
- **Precision** – This criterion examines the precision and accuracy of the indicator based on the following conditions:
 - The margin of error is reported.
 - The margin of error does not significantly change the indicator.
 - The margin of error is acceptable to facilitate decision-making.
 - Program targets include margin of error.
- **Objectivity** – This criterion examines if the objectivity of the indicator based on the following conditions:
 - Data not subject to political and/or personal manipulation.
 - Independence of data collection, management, and assessment.
 - Impartial review of data collection process.
- **Adequacy** – This criterion is another way of measuring instrument validity, based on the following conditions:
 - The indicator fully measures intended outcome.
 - The indicator is sufficient to characterize the outcome.
- **Practicality** – This criterion examines the practicality of collecting and reporting the instrument, based on:
 - Data collection process is viable.

- Data collection is overly costly compared to the adequacy of the indicator.
- Data collection is feasible.

To be able to assess the quality of the performance indicators based on these criteria, the IMPAQ Team performed a number of activities, including:

- Review all relevant program and M&E documentation (including the M&E Plan), manuals, and records.
- Meetings and teleconferences with: 1) the MCG Management team and its IEs responsible for implementation of Program activities; 2) private companies contracted by MCG; and 3) government entities responsible for data collection, data processing and data analysis.
- Review of methodology for estimation of baselines and targets for existing performance indicators and their adequacy with relevant activity outcomes.
- Review of historical data and survey data sets.
- Observation and review of database content to verify the quality and consistency of data.

The IMPAQ Team also reviewed the spot-check methods and the frequency of their utilization by MCG and provided recommendations for improvement. Additionally, we analyzed both primary and secondary data sources in order to observe and verify how data were actually collected and calculated.

To evaluate the performance indicators based on the 7 criteria (and their sub-criteria, as outlined above), we developed a methodology that takes into account the rating of each indicator based on each sub-criterion. Specifically, we developed a scoring system to classify the importance of each sub-criterion as follows:

- 1: Unimportant** – Does not at all meet the specific definition of the criterion.
- 2: Of little importance** – Does not sufficiently meet the specific definition of the criterion.
- 3: Moderately important** – Somewhat meets the specific definition of the criterion.
- 4: Important** – Generally meets the specific definition of the criterion.
- 5: Highly important** – Fully meets the specific definition of the criterion.

The importance of each sub-criterion represents the potential score that an indicator may receive.

For example, Table 2.2 shows the potential scores of the five sub-criteria for the *Validity* criterion. As shown, one highly important factor in assessing the validity of an instrument is whether a relationship between the indicator and the program exists. This criterion has a potential score of 5 for evaluating indicator validity; an indicator can, therefore be rated 1 through 5. If it is rated 5, it indicates that there is a strong relationship between the indicator and the program. A rating of 1 on this sub-criterion indicates a weak relationship. In contrast, whether the indicator has a verifiable source is less critical, so it has a potential score of 2 in assessing instrument validity; thus, an indicator can only be rated 1 through 2 based on this sub-criterion.³ Similarly, we determined the potential scores of the sub-criteria under each of the 7 instrument assessment criteria.

Table 2.2: Instrument Validity Sub-Criteria

	Criteria	Potential Score (1-5)
1	Validity	
1.1	Relationship between indicator and program exists	5
1.2	Indicators defined well and are data reported in an appropriate format	3
1.3	The indicators accurately represent reality	3
1.4	Indicators have a verifiable source	2
1.5	Indicator presents all necessary data	2
	TOTAL	15

Once the score of the indicator is determined based on each sub-criterion, we are able to: 1) construct the overall rating of the indicator for each of the 7 major criteria and 2) construct the total rating for the indicator based on all 7 major criteria. To illustrate how this methodology is implemented, Table 2.3 presents the rating of the *Annual Average Daily Traffic (AADT)* indicator. As indicated in Table 2.3, for sub-criterion 1.1, the actual score is 5, indicating that there is a very close relationship between the indicator and the program.

³ Determining appropriate or adequate thresholds of indicator and data quality is not an exact science. We understand and acknowledge that the items selected to measure each of the seven criteria will display some variation and will be ranked differently based on how well they measure and represent each criterion; however, we have thoroughly reviewed the standard definitions of these criteria and assessed whether they have been used correctly or incorrectly by MCG and its implementing entities as measures of effectiveness for the performance indicators.

Table 2.3: Rating of Average Annual Daily Traffic (AADT) Indicator

	Criteria	Potential Score (1-5)	Actual Score (1-5)	Overall Rating
1	Validity			
1.1	Relationship between indicator and program exists	5	5	
1.2	Indicators defined well and are data reported in an appropriate format	3	3	
1.3	The indicators accurately represent reality	3	3	
1.4	Indicators have a verifiable source	2	2	
1.5	Indicator presents all necessary data	2	2	
	TOTAL	15	15	1.00
2	Reliability			
2.1	Data collected and reported consistently over time, across project activities, and regions	5	5	
2.2	Data collection and reporting process guards against bias	5	4	
2.3	Process is reviewed periodically	4	4	
2.4	Random checks at each stage	4	3	
2.5	Data collection process is documented and problems are reported	2	2	
	TOTAL	20	18	0.90
3	Timeliness			
3.1	Data collected and reported on a frequent basis	5	5	
3.2	Reported data are sufficiently up to date	5	5	
3.3	Date of data collection is clearly identified	5	5	
	TOTAL	15	15	1.00
4	Precision			
4.1	Margin of error is reported	5	5	
4.2	Margin of error does not change the indicator	4	4	
4.3	Margin of error acceptable for decision-making	5	5	
4.4	Program targets include margin of error	4	4	
	TOTAL	18	18	1.00
5	Objectivity			
5.1	Data not subject to political and/or personal manipulation	5	4	
5.2	Independence of data collection, management, and assessment	5	5	
5.3	Impartial review of data collection process	4	4	
	TOTAL	14	13	0.93
6	Adequacy:			
6.1	Indicator fully measures intended outcome	5	4	
6.2	Indicator sufficiently characterizes outcome	5	5	
	TOTAL	10	9	0.90
7	Practicality			
7.1	Data collection process is viable	5	4	
7.2	Data collection not overly costly compared to the adequacy of the indicator	5	4	
7.3	Data collection is feasible	5	5	
	TOTAL	15	13	0.87
OVERALL RATING				6.60

The actual scores for sub-criteria 1.2 through 1.5 were also equal to the potential scores, indicating that all of the validity sub-criteria scored very high, reaching their maximum potential score. Dividing the total actual scores for AADT under the validity criterion (15) by the total potential scores under the validity criterion (15), we find that the overall rating for AADT in terms of validity is 1.00. This indicates AADT received the maximum potential score on all the validity sub-criteria.

AADT also scored high based on other criteria, including reliability (0.90), timeliness (1.00), objectivity (0.93), adequacy (0.90), and practicality (0.87). Adding the overall scores of AADT under each of the 7 criteria, the overall rating of AADT is 6.60 (out of a potential 7.00), indicating that this is a very strong measure for capturing program performance.

2.3 Ratings of Performance Indicators

Using the methodology outlined above, we rated each of the 17 indicators for Round IV. The detailed results of our analysis for each indicator are presented in Appendix A. Table 2.4 summarizes the results of our analysis of all 17 indicators. For each of the 17 indicators, the table presents:

- 1) the total score under each of the 7 criteria, and
- 2) The overall rating based on all 7 criteria.

As shown in Table 2.4, Round IV indicators received very high scores in terms of validity, as all indicators score 0.80 or higher. In fact, the average validity score across all indicators is 0.93 out of a potential 1.00. Similarly, all but one indicator scored 0.80 or higher in reliability, while indicators scored a perfect 1.00 in timeliness and precision. The average objectivity, adequacy, and practicality scores were 0.91 or higher, providing further evidence of the high quality of the indicators. Note that the average scores for the Round IV indicators under each criterion matched or exceeded the respective averages for the Round III indicators, providing evidence that the quality of the data has improved.

Table 2.4: Summary Ratings for Round IV Performance Indicators

Performance Indicator	Rating for Individual Criteria							Overall Rating
	Validity	Reliability	Timeliness	Precision	Objectivity	Adequacy	Practicality	
GOAL INDICATORS								
Poverty gap in the Samtshke-Javakheti Region	0.87	0.95	1.00	1.00	1.00	0.80	0.93	6.55
Poverty Incidence in the Samtshke-Javakheti Region	1.00	0.85	1.00	1.00	1.00	1.00	0.87	6.72
OBJECTIVE INDICATORS								
Jobs created from Enterprise Development Activities	0.87	0.95	1.00	1.00	0.79	0.90	0.93	6.44
Household Net Income	1.00	1.00	1.00	1.00	1.00	1.00	1.00	7.00
OUTCOME INDICATORS, S-J ROAD REHABILITATION								
Savings in Vehicle Operating Costs (VOC)	0.87	0.90	1.00	1.00	0.93	0.90	0.87	6.46
International Roughness Index (IRI)	0.93	0.85	1.00	1.00	0.93	0.90	1.00	6.61
Annual Average Daily Traffic (AADT)	1.00	0.90	1.00	1.00	0.93	0.90	0.87	6.60
Travel Time	0.80	0.75	1.00	1.00	1.00	1.00	1.00	6.55
OUTCOME INDICATORS, GEORGIA REGIONAL DEVELOPMENT FUND								
Increase in gross revenues of Portfolio Companies	0.87	0.90	1.00	1.00	0.86	0.90	1.00	6.52
Increase in PC Employees	0.93	0.80	1.00	1.00	0.93	0.90	0.93	6.50
Increase in local Supplies to the PCs	0.80	0.85	1.00	1.00	1.00	0.80	0.87	6.32
Increase in wages paid to the PC employees	0.93	0.95	1.00	1.00	0.79	0.90	0.87	6.44
Increase in locally sourced goods and services purchased by PCs	1.00	0.90	1.00	1.00	0.86	1.00	0.93	6.69
OUTCOME INDICATORS, AGRIBUSINESS DEVELOPMENT ACTIVITY								
Jobs Created	1.00	0.80	1.00	1.00	0.86	0.80	0.87	6.32
Household Net Income	1.00	0.95	1.00	1.00	1.00	1.00	1.00	6.95
Firm Income	0.87	0.95	1.00	1.00	0.93	0.80	1.00	6.55
Beneficiaries	1.00	1.00	1.00	1.00	1.00	1.00	1.00	7.00
AVERAGE	0.93	0.90	1.00	1.00	0.93	0.91	0.94	6.60

In terms of their overall rating, all indicators scored 6.32 or higher, with an average overall rating of 6.60. This average much exceeded the average overall rating of indicators in previous Rounds. As shown in Table 2.5, the overall rating for Round IV (6.60) much exceeded the rating for 5.36 in Rounds I and II (5.36) and the rating for Round III (6.25). These figures illustrate the steady progress in the overall quality of the indicators over time. A close look at the numbers in Table 2.5 shows that the performance indicators ratings for each individual criterion also improved with each survey.

Based on these results, we conclude that the data quality was very good overall. It is also apparent that the quality of the program indicators has improved from previous Rounds, indicating that the data collection administrators have done a very good job of raising the quality of the data collection process.

Table 2.5: Summary Ratings of Performance Indicators, Round I-IV

Data Round	Rating for Individual Criteria							Overall Rating
	Validity	Reliability	Timeliness	Precision	Objectivity	Adequacy	Practicality	
Round I	0.86	0.78	0.94	0.66	0.71	0.78	0.63	5.36
Round II	0.87	0.79	0.93	0.66	0.71	0.77	0.62	5.36
Round III	0.93	0.86	1.00	1.00	0.78	0.92	0.76	6.25
Round IV	0.93	0.90	1.00	1.00	0.93	0.91	0.94	6.60

CHAPTER 3. REVIEW OF THE SURVEYS

3.1 Introduction

Over the course of MCG's implementation, the IMPAQ Team has conducted annual data quality reviews of six surveys used to collect data for conducting impact evaluations for MCG's two major projects:

- The Regional Infrastructure Rehabilitation (RIR) Project, comprised of the Samtskhe-Javaketi Road Rehabilitation, the Main Gas Pipeline Rehabilitation, and the Regional Infrastructure Development (RID).
- The Enterprise Development Project, comprising of the Georgia Regional Development Fund (GRDF) and the Agribusiness Development Activity (ADA).

These reviews included analyses of the objective of each of the surveys associated with each project, a detailed description of the sampling designs, an analysis of the quality of the questionnaires, training manuals, instructions for filling questionnaires, incentive strategy, quality control methods, data sets, data cleaning rules, etc. These reviews have generated meaningful recommendations of how MCG could build upon their successful data collection strategies while improving in other areas contributing to data inconsistencies.

During Round IV of the DQR, the IMPAQ Team conducted a review of five key surveys. Where possible, the data were compared to Round III to assess any changes in data quality between rounds. The surveys being reviewed for Round IV are the following:

- ADA Beneficiary Survey
- Settlement Infrastructure Survey
- Road Users Survey
- Village Infrastructure Census
- Integrated Household Survey

This section describes the methodology used to assess the quality of these data, provides an assessment of the data quality, and compares the quality of Round IV data to the quality of data received in previous rounds. Section 3.2 provides an assessment of the quality of the data based

on the 7 criteria used to evaluate the key program indicators in Chapter 2. Section 3.3 presents an assessment of the data quality in the ADA Beneficiary Survey; Section 3.4 presents an assessment of the Settlement Infrastructure Survey (SIS); Sections 3.5 – 3.7 will contain assessments of the Road Users Survey, the Village Infrastructure Census (VIC), and the Integrated Household Survey (HIS), respectively.

3.2 Survey Quality Ratings

To assess the overall quality of the Round IV data, the IMPAQ Team employed the same methodology used to assess the quality of the 17 program indicators, as described in Section 2.2. Using a similar approach, we assessed the quality of the Round IV data based on seven criteria: validity, reliability, timeliness, precision, objectivity, adequacy, and practicality. This assessment is critical in understanding the quality of the data and the degree to which they can be used to monitor and evaluate program performance.

Table 3.1 provides a summary of the quality assessment of each Round IV survey, including: 1) the total score under each of the 7 assessment criteria and 2) the overall rating based on all 7 criteria. (Appendix B provides a detailed breakdown of the ratings of each survey based on all sub-criteria.) As shown in Table 3.1, the ADA Beneficiary Survey scored a maximum rating of 1.00 in 4 out of 7 criteria, while it scored at least a 0.93 rating in the remaining criteria. As a result, this survey received an overall rating of 6.81 out of a possible 7.00, which indicates that the information collected and reported by the ADA Beneficiary Survey is a highly dependable source of information for monitoring and assessing program performance.

The Village Infrastructure Census (VIC) and the Integrated Household Survey (IHS) data received maximum rating in all except one criterion and both surveys scored an overall rating of 6.93. Similarly, the overall rating of the Settlement Infrastructure Survey (SIS) and the Road Users Survey were very high, but slightly lower than the overall ratings of the other three data. As a result, the average overall rating of the Round IV data was 6.83 out of 7.00, indicating the high quality of the data.

Table 3.1: Summary Ratings for Round IV Data

Round IV Data	Rating for Individual Criteria							Overall Rating
	Validity	Reliability	Timeliness	Precision	Objectivity	Adequacy	Practicality	
ADA Beneficiary Survey	1.00	1.00	0.93	0.95	0.93	1.00	1.00	6.81
Settlement Infrastructure Survey	1.00	1.00	0.80	1.00	1.00	1.00	0.93	6.73
Road Users Survey	1.00	1.00	0.93	0.95	0.93	1.00	0.93	6.74
Village Infrastructure Census (VIC)	1.00	1.00	1.00	1.00	1.00	1.00	0.93	6.93
Integrated Household Survey (IHS)	1.00	1.00	1.00	1.00	0.93	1.00	1.00	6.93
AVERAGE	1.00	1.00	0.93	0.98	0.96	1.00	0.96	6.83

Furthermore, Table 3.2 provides a summary of the overall rating for each survey in each of the four Rounds, as available. As shown in Table 3.2, the overall quality of the data in Round IV improved relatively to previous Rounds. For example, the overall rating of the ADA Beneficiary Survey in Round IV was 6.81, markedly higher than the 6.55 rating in Round III. Similarly, SIS’s overall rating increased from 6.69 in Round I to 6.73 in Round IV and the Road Survey’s rating increased from 6.65 in Round III to 6.74 in Round IV.

Table 3.2: Overall Ratings of Surveys, Rounds I-IV

Data Source	Overall Rating			
	Round I	Round II	Round III	Round IV
Agribusiness Development Activity (ADA) Beneficiary Survey	-	6.67	6.55	6.81
Settlement Infrastructure Survey (SIS)	6.69	-	-	6.73
Road Users Survey	-	-	6.65	6.74
Village Infrastructure Census (VIC)	-	6.90	-	6.93
Integrated Household Survey (IHS)	-	6.70	6.76	6.93
Regional Infrastructure Development (RID) Project	-	-	6.83	-

These results indicate that data collection administrators did an excellent job of collecting and preparing the information reported in the Round IV data. It is also apparent that they have adopted the recommendations made in the reviews of earlier Rounds of data; they have improved the data collection, coding and documentation strategies; and invested effectively in training

interviewers, coders, and supervisors. Based on these analyses, we conclude that the overall quality of the Round IV data was extremely high overall. The high overall ratings achieved by all surveys indicate that the information reported is highly dependable and program administrators can use it to effectively monitor program performance. Furthermore, the quality of the data suggests that they can be used in the future to potentially evaluate overall program effectiveness.

3.3 ADA Beneficiary Survey

3.3.1 Overview of the ADA Beneficiary Survey

The goal of the Agribusiness Development Activity (ADA) project is to reduce poverty in the regions of Georgia through stimulating economic growth in the agricultural sector. One of the ways that the ADA project accomplishes this goal is through the provision of small and medium-sized matching grant awards, ranging from \$5,000 to \$100,000 (the Enterprise and Value-Chain initiatives). The ADA Beneficiary Survey was developed by MCG to collect data on the program beneficiaries and support the impact evaluations of the project.

There were three components to the Round IV ADA Beneficiary Survey: the Primary Producer Survey, Value Adders- Value Chain Initiative, and the Farm Service Centers. Table 3.3 presents an overview of the sample of these surveys for Rounds III and IV. As shown, the Round IV surveys achieved higher or similar response rates relative to the Round III surveys. We also see, however, that the total sample size was significantly reduced from Round III to Round IV. This is a result of differences in the sampling frame used; Round III data represent cumulative data starting from the baseline data collection and spreading over the period of time ADA grant applications were accepted. Furthermore, Round III data included all individuals who submitted grant applications, including those who submitted more than one application; Round IV data excluded these individuals. Below, we provide a more detailed overview of the quality of the Round IV ADA survey data and how it compares to the quality of the Round III data.

Table 3.3: Sample Characteristics, ADA Surveys

	Survey Name		
	Primary Producer	Value Adders- Value Chain Initiative	Farm Service Centers
Round III			
Sample	527	227	99
Interviewed (response rate)	440 (83%)	152 (67%)	77 (78%)
Round IV			
Sample	252	56	27
Interviewed (response rate)	201 (80%)*	44(79%)*	24(88%)*

Note: *Reported are total respondents with the response rate in parenthesis.

3.3.2 Brief Description of Findings from ADA Surveys

To examine the quality of the ADA surveys, we assess a number of key variables, including gender, region, experimental evaluation group, and the ability to take an agricultural loan. These variables were chosen because they represent key variables of interest in assessing the program population and performance. Moreover, these variables are available in both Round III and Round IV data, enabling us to produce a meaningful comparison of data quality over time.

Tables 3.4 through 3.6 present the gender distribution for the three ADA surveys in Rounds III and IV. The results indicate that the gender distribution in all three surveys remained relatively constant between Round III and Round IV.

Table 3.4: Gender Distribution, Primary Producer Survey Respondents

Values	Primary Producer Survey	
	Round III	Round IV
Male	349 (79%)	151 (75%)
Female	25 (6%)	12 (6%)
Missing	66 (15%)	38 (19%)
Values not Labeled	-	-
Total Number of Respondents	440 (100%)	201 (100%)

Note: Reported is the number of respondents with proportion of total respondents in parenthesis.

Table 3.5: Gender Distribution, Value Adders- Value Chain Survey Respondents

Values	Value Adders- Value Chain Survey	
	Round III	Round IV
Male	116 (76%)	30(68%)
Female	17 (11%)	4 (9%)
Missing	19 (13%)	10 (22%)
Values not Labeled	-	-
Total Number of Respondents	152(100%)	44 (100%)

Note: Reported is the number of respondents with proportion of total respondents in parenthesis.

Table 3.6: Gender Distribution, Farm Service Center Survey Respondents

Values	Farm Service Center Survey	
	Round III	Round IV
Male	64 (83%)	17 (71%)
Female	6 (8%)	4 (17%)
Missing	7 (9%)	3 (12%)
Values not Labeled	-	-
Total Number of Respondents	77 (100%)	24 (100%)

Note: Reported is the number of respondents with proportion of total respondents in parenthesis.

A close examination of Tables 3.4 through 3.6 reveals a substantial number of missing data values in Rounds III and Round IV. It is reassuring to note that these missing data values do not appear to be a problem with coding of the surveys. That is, individuals whose businesses had ceased operations or merged with other businesses were administered a different and much shorter survey. The short telephone interview that was conducted with these respondents did not include many variables of interest, including gender. Thus, the problem is not with the data collection or coding, but rather with the original design of the survey instrument for these individuals. To deal with this issue, it may be helpful to assign a flag to all respondents that completed the shorter survey so that data users could appropriately account for the non-responses. Accounting for these entities (i.e., entities that merged or ceased to operate) leads to a zero rate of missing data for Round IV.

Missing rates for Round III on the other hand reflect survey respondents who were already interviewed on the basis of another grant application. Accounting for these respondents also brings the missing data rate to zero. While we are now able to understand why this (and other variables) has numerous missing data values in for each round, other researchers may have difficulties in figuring this out from the dataset or its documentation. It may be helpful to create a flag that notes these reasons for “missing data.” While we used one variable to illustrate this issue, other variables are susceptible to the same “missing data” issues.

Tables 3.7 through 3.9 highlight a relatively minor issue identified in Round III that appears to have been reduced in Round IV across the three ADA surveys. The IMPAQ Team has recently learned that this is not a data entry issue at all. Specifically, the ‘region’ variable was not data entered by IPM; rather, this variable is simply taken as given by CNFA. Thus any “improvement” is due to the different respondent samples in Round III and Round IV.

Table 3.7: Regional Distribution, Primary Producer Survey Respondents

Regions	Primary Producer Survey	
	Round III	Round IV
Adjara	14	6
Guria	15	4
Imereti	58	19
Kakheti	94	49
Mtskheta-Tianeti	27	11
Kvemo Kartli	81	24
Racha-Lechkumi and Kvemo Svaneti	7	1
Samtskhe-Javakheti	42	8
Samegro-Zemo Svaneti	-	14
Shida Kartli	66	27
Samegrelo	33	-
Svaneti	3	-
Administered shorter interviews	-	38*
Values not Labeled	-	-
Total Number of Respondents	440 (100%)	201 (100%)

Note: *The shorter interview did not include the region question.

Table 3.8: Regional Distribution, Value Adders- Value Chain Survey Respondents

Regions	Value Adders- Value Chain Survey	
	Round III	Round IV
Adjara	14	6
Guria	17	4
Imereti	19	6
Kakheti	39	7
Mtskheta-Tianeti	18	1
Kvemo Kartli	5	2
Racha-Lechkumi and Kvemo Svaneti	3	1
Samtskhe-javakheti	12	2
Samegro-Zemo Svaneti	-	3
Shida Kartli	7	2
Samegrelo	18	-
Svaneti	-	-
Administered shorter interviews	-	10*
Values not Labeled	-	-
Total Number of Respondents	152 (100%)	44 (100%)

*Note: The shorter interview did not include the region question.

Table 3.9: Regional Distribution, Farm Service Center Survey Respondents

Regions	Farm Service Center Survey	
	Round III	Round IV
Adjara	3	2
Guria	5	1
Imereti	15	6
Kakheti	17	8
Mtskheta-Tianeti	2	-
Kvemo Kartli	5	3
Racha-Lechkumi and Kvemo Svaneti	-	-
Samtskhe-javakheti	11	1
Samegro-Zemo Svaneti	-	-
Shida Kartli	8	3
Samegrelo	10	-
Svaneti	1	-
Administered shorter interviews	3	2*
Values not Labeled	5	1
Total Number of Respondents	77 (100%)	24 (100%)

* Note: The shorter interview did not include the region question.

Another infrequent issue is illustrated in Tables 3.10 through 3.12. Taking the variable ‘Ability to take a loan’ as an example across the three ADA surveys, we note that not all values for the variable are labeled. It is difficult to determine whether the three unlabeled values in Round III of the Value Adders-Value Chain Survey or the one value in Round IV of the same survey are missing values or a valid skip. As discussed above for the ‘region’ variable, the ‘Ability to take a loan’ is also provided by CNFA without any editing or validation. Thus, any change between rounds is due to a difference in the respondents at each round.

Table 3.10: Ability to take a loan, Primary Producer Survey Respondents

Values	Primary Producer Survey	
	Round III	Round IV
Valid	439 (100%)	200 (100%)
Missing	1 (0%)	1 (0%)
Values not Labeled	-	-
Total Number of Respondents	440 (100%)	201 (100%)

Note: The missing value is recorded as “N/A” in the survey for both rounds.

Table 3.11: Ability to take a loan, Value Adders-Value Chain Survey Respondents

Values	Value Adders- Value Chain Survey	
	Round III	Round IV
Valid	149 (98%)	43 (98%)
Missing	-	-
Values not Labeled	3 (2%)	1 (2%)
Total Number of Respondents	152 (100%)	44 (100%)

Note: The missing value is recorded as “N/A” in the survey for both rounds.

Table 3.12: Ability to take a loan, Farm Service Center Survey Respondents

Values	Farm Service Center Survey	
	Round III	Round IV
Valid	75 (97%)	24 (100%)
Missing	-	-
Values not Labeled	2 (3%)	-
Total Number of Respondents	77 (100%)	24 (100%)

Note: The missing value is recorded as “N/A” in the survey for both rounds.

3.3.3 General Comments on the ADA Survey

The field work procedures used by IPM were designed and planned in previous rounds of the DQR. The procedures were carefully reviewed in previous rounds and improvements have been made. Due to high non-response rates for some key variables in previous rounds, IPM decided to redesign some of the questions in the interview. To redesign the interview, IPM conducted a

number of cognitive interviews and consultation with experts. The new design was pilot tested to assess the improvements. Following the pilot test, IPM developed enhanced training materials to train interviewers in the use of the new instrument. As a result of these changes to the questionnaire, response rates to key questions increased, as has been documented by IPM. In addition, the use of 100% double data entry has also minimized data entry errors. Overall, the quality of the data collected has improved and the issues highlighted above are minor and would not impact the quality of monitoring and evaluation. As the improvement of the survey can be attributed to an iterative process of data quality reviews and applying recommendations, the steps that have been taken over the previous rounds of the DQR are described below.

Through Round II, the ADA data entry was done by only one operator with many years of experience. Because each ADA survey has a small sample, only one operator keyed the data. No data quality assurance processes were in place for data entry. IMPAQ recommended 15% double entry and calculation of error rate as a first step, with periodic monitoring, to establish and maintain a rate within industry-accepted standards.

Following IMPAQ's recommendation regarding double keying, IPM first calculated the error rates and found them to be 1.7% (figured as number of keystrokes in error over total keystrokes for an instrument) for the FSC, 2.9% for the VA/VCI, and 2.7% for the PP surveys. As a result of these findings, it was reported to IMPAQ that the ADA is now keyed by two operators, and 100% of the data are double keyed. IMPAQ agrees that 100% double entry is the most suitable approach for surveys with so few observations, where even a few errors may have major impact on indicators.

Overall, the quality of the ADA Beneficiary Survey is very high, making it a reliable data source for monitoring and evaluating program performance. These data were rated very high based on the seven criteria used to assess data quality (see Table 3.1); as a result, the overall rating of the ADA data was 6.81 out of a possible 7.00. The high rating illustrates the quality of the data and highlights the improvement in data quality achieved by data collectors.

3.4 Settlements Infrastructure Survey (SIS)

3.4.1 Overview of the SIS

One of the activities of the RIR project is the Samtskhe-Javakheti Road Rehabilitation. The Settlements Infrastructure Survey (SIS) is a survey produced for monitoring the performance of this project. SIS is conducted in settlements (i.e., villages, towns, or cities) located in Tbilisi, as well as in all the local administrations and district centers of five districts located along Samtskhe-Javakheti road.

SIS was conducted twice, with the first survey conducted in 2007 and managed by IPM and the second survey conducted in 2010 by ACT Research. SIS collects 16 different types of individual, infrastructure, and economic environment information, including data on respondents, geographic information, utilities infrastructure, transportation, agriculture, prices, industry and construction, healthcare, and tourism.

3.4.2 Brief Description of Findings from SIS Round IV Survey (2010)

The IMPAQ Team's review of the Round IV SIS data showed that the quality of data collected is very high. This was achieved despite the addition and modification of some questions from the Baseline SIS Survey for administration in Round IV. The high quality of the data is reflected in an overall rating score of 6.73 out of a possible 7.0 (see Table 3.1). This can be attributed to a great investment in data quality controls during various stages of the data collection process, from pilot testing new question formats to undertaking a three-step approach to data entry quality controls. A random sampling of approximately 15 variables revealed no major data collection or entry inconsistencies. The rates for missing data are within an acceptable range of five percent or less.

3.4.3 General Comments on SIS

Following IMPAQ's Round I review and analysis of the SIS instrument and supporting documents, IMPAQ identified a number of issues and areas for improvements of the SIS. For example, the survey instrument required the respondent to provide great deal of information which may not be easily available. In some cases, providing accurate responses to a number of these questions required consulting records. This presented an undue burden on respondents and

in other cases records were not easily available or required respondents to recall information from their past. Recalling information from the distant past is susceptible to imprecise estimates by respondents. In all, these issues presented challenges to the accurate collection of information.

To address these issues, among others, the SIS Baseline instrument was revised for its administration in Round IV. Some questions were dropped, others had revised timelines, and some new questions were added. In addition, a fairly detailed field interviewer training process was instituted at the primary data control level. Interviewers were trained in collecting information from group discussions, rules for filling out questionnaires, and recruitment procedures.

Overall, the training plan and the actual description of training are very good. The training of enumerators and others to administer the SIS was also of high quality. Generally, the procedures are very thorough and in line with best practices in training of survey interviewers. Training sessions included a thorough discussion of project goals as well as motivations for the questions in the instruments. Training also covered skip instructions, how to record answers, and how to probe answers such as “don’t know” or “refused”. In addition, the training sessions included detailed training on how to approach each settlement, how to set up appointments, etc.

We conclude that the information collected by SIS in Round IV was of very high quality, making SIS a reliable measure of program performance. This is to a great extent a result of the high quality of training provided to SIS enumerators and other survey administrators. We do not have any additional suggestions for improving the quality of the survey and we expect that future data collection efforts will be of very high quality.

3.5 Road Users Survey

3.5.1 Overview of the Road Survey

The Road Survey is a group of surveys conducted by GORBI focusing on information collection regarding road usage. There are three waves of data collection, each wave consisting of a one-week period (seven consecutive days). Three distinct surveys were conducted: 1) the Traffic Count Survey, 2) the Speed Measurement Survey, and 3) the Origin-Destination Survey.

3.5.2 Brief Description of Findings from the Road Survey

The IMPAQ Team reviewed the information provided by each survey in the Round IV Road Survey data to assess the overall quality of the data. During the review, no major issues were identified, indicating that the overall quality of the data was very high. Specifically:

- **Traffic Count Survey** – The Traffic Count Survey measures changes in traffic flows through selected observation points and is conducted approximately once every three months. Each record in the data file reported the number of vehicles of the designated vehicle type that passed the designated observation point in the designated direction in each of the allocated time slots. Based on our review of these data, we found no major inconsistencies in the reported information.
- **Speed Measurement Survey** – IMPAQ was unable to access the data files provided for the Speed Measurement category, as they were in a format requiring ArcView to access. MCG should consider converting these data to a format similar to the rest of the data so that the data could be adequately evaluated.
- **Origin-Destination Survey** – This survey provides information about road travel at the Georgia border crossings, including counts of vehicles crossing the border and voluntary and self-administered interviews of drivers crossing the border. The interviews provide a number of variables, including the driver’s purpose of crossing the border and the types of goods being transported on the road. Based on our review of several variables in the data, we did not find any major inconsistencies. The high quality of the data can be attributed to the data quality check processes used, which included a supervisor reviewing the information collected by interviewers; field supervisors conducting periodic site visits; and a cross-validation of the data at the data entry point by a second party.

3.5.3 General Comments on the Road Survey

According to the methodology overview provided to IMPAQ, the data represent points of observation/instances over three one-week observation periods. In addition, data are collected over 24 hour periods. This traffic activity sampling framework provides a highly reliable base for projecting annual road use and traffic patterns. One area of concern is that if data collection is done on days that are not typical (e.g., prolonged holidays and days with inclement weather

conditions), such estimates would be misleading. However, we did not identify such issues in the data reviewed.

In addition to the superior traffic and road use sampling, the data quality control methods used limit the possibility for errors and data inconsistencies. Speed measurement was collected using a combination of video capture and GPS devices to track vehicle trips. In addition, a second enumerator took notes of driving conditions to complement the objective data collected by the electronic methods. A combination of these methods enabled a very close approximation of “normal” speed conditions for each portion of the road assessed. Finally, each trip took two measurements, including the return trip, enabling enumerators to assess the impact of any diversions on their trip.

Adequate steps were also taken to ensure that the origin-destination data were consistently recorded at border crossings. Each data collection shift included one vehicle counter and two interviewers collecting written surveys from drivers. Each shift had back-up reserves to fill in incase an interviewer had to leave the shift prematurely. All these steps ensured that each moment during the 24-hour data collection period was covered. Moreover, the large majority of drivers completed the interviews (659 out of 672).

Finally, the Traffic Count survey mitigated the possibility of collecting unreliable data by having interviewers submit their collection forms to a supervisor who would verify them. The supervisor in turn would submit the data to a home office where data would be verified again. The cumulative contribution of these meticulous steps contributed to the Road Surveys having high validity and reliability scores.

3.6 Village Infrastructure Census (VIC)

3.6.1 Overview of the VIC

The main purpose of the VIC survey is to collect information on the infrastructure available to the villages, including the availability of gas supply, electricity, drinking water, irrigation systems, and other services. In addition, the VIC provides information on the existence of infrastructure projects at the village level, including: health care, education facilities, banks,

processing agricultural enterprises, delivery facilities, and other objects. The first VIC data collection took place in 2008 and data were collected from all communities and villages across Georgia. The second VIC data collection started in May 2010 based on an enhanced data collection methodology. Specifically, interviewers introduced the use of voice recorders to augment hand-written completion of survey questionnaires. This study has been conducted specifically for MCG by the Department of Statistics (DS). As the primary sponsor of the study, MCG has had a great deal of input as to how this study has been done. Nonetheless, in working with the national statistical agency, MCG has had to allow the DS to follow some of its own standard operating procedures.

3.6.2 Brief Description of Findings from VIC

Overall, the VIC did not have major issues with variable inconsistencies. To examine the quality of the VIC, we examined 30 randomly selected variables. Based on our review of these variables, no major data quality issues were observed, largely due to the comprehensive steps taken to train interviewers and to establish numerous data quality control procedures.

One indicative example of the improvement in the quality of the data is the improvement in the process used to produce and use skip patterns to minimize the potential for errors in counting sample sizes for each question. In Round II, for example, there were cases where skip patterns had to be used to indicate whether individuals would respond to a series of follow-up questions based on their responses to a specific question. However, respondents who were not supposed to respond to certain questions (i.e., those with a skip pattern) were asked by interviewers to respond, leading to misrepresentative results. This may be due to: 1) the lack of appropriate skip patterns coding and/or 2) the interviewers did not accurately observe the skip patterns.

IMPAQ staff observed this issue and made recommendations for correcting it. Indeed, our review of Round IV VIC data indicates that there were no skip pattern issues, suggesting that data administrators successfully implemented and monitored the creation and use of appropriate skip patterns.

3.6.3 General Comments on VIC

The high quality of data collected can be attributed to three measures; the successful completion of a pilot test; comprehensive training of field staff; and a meticulous logical control and data entry phase, enhanced by the use of recorded field interviews. The purpose of the pilot testing was to refine the census tool as well as the data collection methods. To achieve these goals, pilot testing was conducted using identical procedures to be applied with the actual VIC survey. These procedures included the selection of villages, selection of field staff, training of field staff, and recruitment of respondents. Pilot testing enabled administrators to recognize and provide solutions to several problems including:

- *Village status* – The status or classification of villages had undergone some changes since the first round of the VIC survey. A decision was taken use the new status, District Municipality instead of Temi. Appropriate steps were taken to ensure District Municipalities were informed about the upcoming VIC survey.
- *Difficulty in reaching all villages due to poor transportation means* – It was decided that in such cases the survey should be conducted in Temi centers.
- *Respondent discrepancy in identifying village borders* – At issue is the assignment of village infrastructure to the appropriate village as some respondents did not know the extent of their villages. To resolve this, it was decided that Temi representatives had to be included in the survey sample.

Selection and training of staff took several phases designed to ensure a high quality data collection. Potential interviewers were first sourced from interviewers with an excellent track record from the first round, followed by experienced GeoStat employees, and then recruits with both survey administration and group interview experience. These recruits were then trained and tested before participating in the survey. Several groups of interviewers were retrained to bring their ability up the level of other interviewers.

Finally, once data was collected in the field, the integrity of the data was verified by trained coders during the logical control and data entry phase. During this phase issues such as mechanical errors, inadvertent question completion when respondents should have been skipped,

inconsistent numbers given for village distance to amenities, and problems with measurement units were addressed. These errors were corrected using the recorded interviews and/or follow-up telephone calls. Data entry errors were limited by using software that controlled for skip logic and coding of responses to appropriate classification nomenclature (e.g. assigning agricultural produce codes).

3.7 Integrated Household Survey (IHS)

3.7.1 Overview of the IHS

The primary purpose of the IHS survey is to support all components of the two major projects of MCG, the Regional Infrastructure Rehabilitation Project and the Enterprise Development Project. This is achieved by enhancing the traditional Household Survey with data on infrastructure, markets, the business environment, and agricultural production. Data are collected from the household member with the most information on household income and expenditures. IHS is conducted by GeoStat at the Ministry of the Economic Department of Georgia on a quarterly basis and included selected households in Tbilisi and in all Georgia regions. The sampling frame for the survey is based on the 2002 population census enumeration units, while the sampling design used a multi-stage cluster approach.

For the second administration of the survey the survey sample was doubled by the GeoStat in order to: 1) update the sampling frame to account for growth and migration and 2) collect more comprehensive information at the household level on infrastructure, markets, and agricultural production. This would enable improved estimates at the regional level.

The IHS questionnaire consists of eight modules collecting information on a different topic:

- Shinda 01 – Data on living conditions and land usage
- Shinda 02 – Data on household composition and its changes
- Shinda 03 – Diary of Expenses (for keeping records on the expenses made by the family members on food and non-food goods and services during a week)
- Shinda 04 – Data on consumption and investment expenses (e.g., housing, land, transport, healthcare, education and child care, ceremonies and land)
- Shinda 05 – Data on transfers and incomes from property, savings, etc.

- Shinda 05_1 – Data on employment and income from employment
- Shinda 07 – (Refusal Sheet) – Information on reasons for non-response
- Shinda 09 – Data on availability of education and healthcare services

3.7.2 Brief Description of Findings from IHS

A review of 50 randomly selected variables from multiple datasets from IHS revealed few negligible errors of low consequence. For example, in some cases, the data entry included values outside the expected range for the specific question. For example respondents were asked to keep a diary of daily expenses on household expenditures. Respondents had to record which day of the week their records corresponded to hence it was not possible to have data recorded for an “8th” day of the week. In this case the percentage of respondents having an “8th” day of the week was less than one percent. Other such errors included one record of having “0” level of education when the minimum possible value was “1”. Such errors would be fairly obvious to dataset users and would have little bearing on final analyses.

3.7.3 General Comments on IHS

Overall, the quality of the IHS was very good and there were few errors. This contributed to the IHS sharing the highest overall quality rating with the VIC survey with a score of 6.93 out of a possible 7. The IHS is a complex survey to administer and requires interviewers to pay close attention to the logical sequence of questions. A review of the training documentation provided shows that training of interviewers included considerable attention to this issue. In addition, the IHS contains numerous retrospective questions that ask respondents to recall items from up to three months prior to the date of interview. Interview trainers made sure to provide a common definition for all interviewers for what “prior three months” consisted of. For example, trainers worked with interviewers to provide examples of household expenses that were to be included for the prior three months. In addition, interviewers were taught techniques to enhance the reliability of data that respondents had to recall. Finally, as the survey asked questions on fairly complex topics, we noted that some questions seemed to be worded awkwardly- perhaps this may be attributed to the challenges of translating complex issues between languages.

CHAPTER 4. SUMMARY AND CONCLUSIONS

The results of this report indicate the high quality of the data that have been collected in the course of implementing the Millennium Challenge Corporation Compact in Georgia. We have found the data to be of sufficient quality to monitor and evaluate the effectiveness of the programs that have been implemented over the course of the compact. Moreover, we have found that the MCA-Georgia M&E team has been a willing partner in improving the data that is being collected in support of the monitoring and evaluation activities.

We developed an evaluation criterion to measure the quality of the data collected and used in each of the indicators and the surveys. In this evaluation criterion, a perfect score is 7. As seen in Table 4.1, in Round IV, each of the surveys had an overall score close to 7, indicating that the data was of high quality.

Table 4.1: Summary Ratings for Round IV Data

Round IV Data	Overall Rating
ADA Beneficiary Survey	6.81
Settlement Infrastructure Survey	6.73
Road Users Survey	6.74
Village Infrastructure Census (VIC)	6.93
Integrated Household Survey (IHS)	6.93
AVERAGE	6.83

Another indication of the efforts made to obtain high quality data is indicated in Table 4.2 where the scores are shown for previous rounds of data collection. As indicated, there has been a steady increase in scores over the years.

Table 4.2: Overall Ratings of Surveys, Rounds I-IV

Data Source	Overall Rating			
	Round I	Round II	Round III	Round IV
Agribusiness Development Activity (ADA) Beneficiary Survey	-	6.67	6.55	6.81
Settlement Infrastructure Survey (SIS)	6.69	-	-	6.73
Road Users Survey	-	-	6.65	6.74
Village Infrastructure Census (VIC)	-	6.90	-	6.93
Integrated Household Survey (IHS)	-	6.70	6.76	6.93
Regional Infrastructure Development (RID) Project	-	-	6.83	-

Our overall conclusion from this study is that the data currently available to assess the MCC Georgia compact is of sufficient quality to meet all the needs of a rigorous evaluation. Over the past few years, MCA-Georgia’s data collection and reporting procedures have substantially improved and the data has the accuracy and reliability needed to rigorously evaluate the MCC Georgia Compact program.

APPENDIX A: RATING TABLES FOR PROGRAM PERFORMANCE INDICATORS

Table A.1: Rating of Poverty Gap in the Samtshke-Javakheti Region Indicator

	Criteria	Potential Score (1-5)	Actual Score (1-5)	Overall Rating
1	Validity			
1.1	Relationship between indicator and program exists	5	4	
1.2	Indicators defined well and are data reported in an appropriate format	3	3	
1.3	The indicators accurately represent reality	3	2	
1.4	Indicators have a verifiable source	2	2	
1.5	Indicator presents all necessary data	2	2	
	TOTAL	15	13	
2	Reliability			
2.1	Data collected and reported consistently over time, across project activities, and regions	5	4	
2.2	Data collection and reporting process guards against bias	5	5	
2.3	Process is reviewed periodically	4	4	
2.4	Random checks at each stage	4	4	
2.5	Data collection process is documented and problems are reported	2	2	
	TOTAL	20	19	
3	Timeliness			
3.1	Data collected and reported on a frequent basis	5	5	
3.2	Reported data are sufficiently up to date	5	5	
3.3	Date of data collection is clearly identified	5	5	
	TOTAL	15	15	
4	Precision			
4.1	Margin of error is reported	5	5	
4.2	Margin of error does not change the indicator	4	4	
4.3	Margin of error acceptable for decision-making	5	5	
4.4	Program targets include margin of error	4	4	
	TOTAL	18	18	
5	Objectivity			
5.1	Data not subject to political and/or personal manipulation	5	5	
5.2	Independence of data collection, management, and assessment	5	5	
5.3	Impartial review of data collection process	4	4	
	TOTAL	14	14	
6	Adequacy:			
6.1	Indicator fully measures intended outcome	5	4	
6.2	Indicator sufficiently characterizes outcome	5	4	
	TOTAL	10	8	
7	Practicality			
7.1	Data collection process is viable	5	5	
7.2	Data collection not overly costly compared to the adequacy of the indicator	5	4	
7.3	Data collection is feasible	5	5	
	TOTAL	15	13	
OVERALL RATING				6.55

Table A.2: Rating of Poverty Incidence in the Samtshke-Javakheti Region Indicator

	Criteria	Potential Score (1-5)	Actual Score (1-5)	Overall Rating
1	Validity			
1.1	Relationship between indicator and program exists	5	5	
1.2	Indicators defined well and are data reported in an appropriate format	3	3	
1.3	The indicators accurately represent reality	3	3	
1.4	Indicators have a verifiable source	2	2	
1.5	Indicator presents all necessary data	2	2	
	TOTAL	15	15	1.00
2	Reliability			
2.1	Data collected and reported consistently over time, across project activities, and regions	5	3	
2.2	Data collection and reporting process guards against bias	5	5	
2.3	Process is reviewed periodically	4	4	
2.4	Random checks at each stage	4	3	
2.5	Data collection process is documented and problems are reported	2	2	
	TOTAL	20	17	0.85
3	Timeliness			
3.1	Data collected and reported on a frequent basis	5	5	
3.2	Reported data are sufficiently up to date	5	5	
3.3	Date of data collection is clearly identified	5	5	
	TOTAL	15	15	1.00
4	Precision			
4.1	Margin of error is reported	5	5	
4.2	Margin of error does not change the indicator	4	4	
4.3	Margin of error acceptable for decision-making	5	5	
4.4	Program targets include margin of error	4	4	
	TOTAL	18	18	1.00
5	Objectivity			
5.1	Data not subject to political and/or personal manipulation	5	5	
5.2	Independence of data collection, management, and assessment	5	5	
5.3	Impartial review of data collection process	4	4	
	TOTAL	14	13	1.00
6	Adequacy:			
6.1	Indicator fully measures intended outcome	5	5	
6.2	Indicator sufficiently characterizes outcome	5	5	
	TOTAL	10	9	1.00
7	Practicality			
7.1	Data collection process is viable	5	4	
7.2	Data collection not overly costly compared to the adequacy of the indicator	5	4	
7.3	Data collection is feasible	5	5	
	TOTAL	15	13	0.87
OVERALL RATING				6.72

Table A.3: Rating of Jobs Created from Enterprise Development Activities Indicator

	Criteria	Potential Score (1-5)	Actual Score (1-5)	Overall Rating
1	Validity			
1.1	Relationship between indicator and program exists	5	5	
1.2	Indicators defined well and are data reported in an appropriate format	3	3	
1.3	The indicators accurately represent reality	3	2	
1.4	Indicators have a verifiable source	2	2	
1.5	Indicator presents all necessary data	2	1	
	TOTAL	15	13	0.87
2	Reliability			
2.1	Data collected and reported consistently over time, across project activities, and regions	5	5	
2.2	Data collection and reporting process guards against bias	5	4	
2.3	Process is reviewed periodically	4	4	
2.4	Random checks at each stage	4	4	
2.5	Data collection process is documented and problems are reported	2	2	
	TOTAL	20	19	0.95
3	Timeliness			
3.1	Data collected and reported on a frequent basis	5	5	
3.2	Reported data are sufficiently up to date	5	5	
3.3	Date of data collection is clearly identified	5	5	
	TOTAL	15	15	1.00
4	Precision			
4.1	Margin of error is reported	5	5	
4.2	Margin of error does not change the indicator	4	4	
4.3	Margin of error acceptable for decision-making	5	5	
4.4	Program targets include margin of error	4	4	
	TOTAL	18	18	1.00
5	Objectivity			
5.1	Data not subject to political and/or personal manipulation	5	5	
5.2	Independence of data collection, management, and assessment	5	4	
5.3	Impartial review of data collection process	4	2	
	TOTAL	14	11	0.79
6	Adequacy:			
6.1	Indicator fully measures intended outcome	5	5	
6.2	Indicator sufficiently characterizes outcome	5	4	
	TOTAL	10	9	0.90
7	Practicality			
7.1	Data collection process is viable	5	5	
7.2	Data collection not overly costly compared to the adequacy of the indicator	5	4	
7.3	Data collection is feasible	5	5	
	TOTAL	15	13	0.93
OVERALL RATING				6.44

Table A.4: Rating of Household Net Income Indicator

	Criteria	Potential Score (1-5)	Actual Score (1-5)	Overall Rating
1	Validity			
1.1	Relationship between indicator and program exists	5	5	
1.2	Indicators defined well and are data reported in an appropriate format	3	3	
1.3	The indicators accurately represent reality	3	3	
1.4	Indicators have a verifiable source	2	2	
1.5	Indicator presents all necessary data	2	2	
	TOTAL	15	15	1.00
2	Reliability			
2.1	Data collected and reported consistently over time, across project activities, and regions	5	5	
2.2	Data collection and reporting process guards against bias	5	5	
2.3	Process is reviewed periodically	4	4	
2.4	Random checks at each stage	4	4	
2.5	Data collection process is documented and problems are reported	2	2	
	TOTAL	20	20	1.00
3	Timeliness			
3.1	Data collected and reported on a frequent basis	5	5	
3.2	Reported data are sufficiently up to date	5	5	
3.3	Date of data collection is clearly identified	5	5	
	TOTAL	15	15	1.00
4	Precision			
4.1	Margin of error is reported	5	5	
4.2	Margin of error does not change the indicator	4	4	
4.3	Margin of error acceptable for decision-making	5	5	
4.4	Program targets include margin of error	4	4	
	TOTAL	18	18	1.00
5	Objectivity			
5.1	Data not subject to political and/or personal manipulation	5	5	
5.2	Independence of data collection, management, and assessment	5	5	
5.3	Impartial review of data collection process	4	4	
	TOTAL	14	14	1.00
6	Adequacy:			
6.1	Indicator fully measures intended outcome	5	5	
6.2	Indicator sufficiently characterizes outcome	5	5	
	TOTAL	10	10	1.00
7	Practicality			
7.1	Data collection process is viable	5	5	
7.2	Data collection not overly costly compared to the adequacy of the indicator	5	5	
7.3	Data collection is feasible	5	5	
	TOTAL	15	15	1.00
OVERALL RATING			7.00	

Table A.5: Rating of Savings in Vehicle Operating Costs (VOC) Indicator

	Criteria	Potential Score (1-5)	Actual Score (1-5)	Overall Rating
1	Validity			
1.1	Relationship between indicator and program exists	5	4	
1.2	Indicators defined well and are data reported in an appropriate format	3	3	
1.3	The indicators accurately represent reality	3	3	
1.4	Indicators have a verifiable source	2	2	
1.5	Indicator presents all necessary data	2	1	
	TOTAL	15	13	0.87
2	Reliability			
2.1	Data collected and reported consistently over time, across project activities, and regions	5	4	
2.2	Data collection and reporting process guards against bias	5	4	
2.3	Process is reviewed periodically	4	4	
2.4	Random checks at each stage	4	4	
2.5	Data collection process is documented and problems are reported	2	2	
	TOTAL	20	18	0.90
3	Timeliness			
3.1	Data collected and reported on a frequent basis	5	5	
3.2	Reported data are sufficiently up to date	5	5	
3.3	Date of data collection is clearly identified	5	5	
	TOTAL	15	15	1.00
4	Precision			
4.1	Margin of error is reported	5	5	
4.2	Margin of error does not change the indicator	4	4	
4.3	Margin of error acceptable for decision-making	5	5	
4.4	Program targets include margin of error	4	4	
	TOTAL	18	18	1.00
5	Objectivity			
5.1	Data not subject to political and/or personal manipulation	5	5	
5.2	Independence of data collection, management, and assessment	5	5	
5.3	Impartial review of data collection process	4	3	
	TOTAL	14	13	0.93
6	Adequacy:			
6.1	Indicator fully measures intended outcome	5	5	
6.2	Indicator sufficiently characterizes outcome	5	4	
	TOTAL	10	9	0.90
7	Practicality			
7.1	Data collection process is viable	5	4	
7.2	Data collection not overly costly compared to the adequacy of the indicator	5	5	
7.3	Data collection is feasible	5	4	
	TOTAL	15	13	0.87
OVERALL RATING				6.46

Table A.6: Rating of International Roughness Index (IRI) Indicator

	Criteria	Potential Score (1-5)	Actual Score (1-5)	Overall Rating
1	Validity			
1.1	Relationship between indicator and program exists	5	5	
1.2	Indicators defined well and are data reported in an appropriate format	3	3	
1.3	The indicators accurately represent reality	3	2	
1.4	Indicators have a verifiable source	2	2	
1.5	Indicator presents all necessary data	2	2	
	TOTAL	15	14	
2	Reliability			
2.1	Data collected and reported consistently over time, across project activities, and regions	5	4	
2.2	Data collection and reporting process guards against bias	5	4	
2.3	Process is reviewed periodically	4	4	
2.4	Random checks at each stage	4	3	
2.5	Data collection process is documented and problems are reported	2	2	
	TOTAL	20	17	
3	Timeliness			
3.1	Data collected and reported on a frequent basis	5	5	
3.2	Reported data are sufficiently up to date	5	5	
3.3	Date of data collection is clearly identified	5	5	
	TOTAL	15	15	
4	Precision			
4.1	Margin of error is reported	5	5	
4.2	Margin of error does not change the indicator	4	4	
4.3	Margin of error acceptable for decision-making	5	5	
4.4	Program targets include margin of error	4	4	
	TOTAL	18	18	
5	Objectivity			
5.1	Data not subject to political and/or personal manipulation	5	5	
5.2	Independence of data collection, management, and assessment	5	4	
5.3	Impartial review of data collection process	4	4	
	TOTAL	14	13	
6	Adequacy:			
6.1	Indicator fully measures intended outcome	5	5	
6.2	Indicator sufficiently characterizes outcome	5	4	
	TOTAL	10	9	
7	Practicality			
7.1	Data collection process is viable	5	5	
7.2	Data collection not overly costly compared to the adequacy of the indicator	5	5	
7.3	Data collection is feasible	5	5	
	TOTAL	15	15	
OVERALL RATING				6.61

Table A.7: Rating of Annual Average Daily Traffic (AADT) Indicator

	Criteria	Potential Score (1-5)	Actual Score (1-5)	Overall Rating
1	Validity			
1.1	Relationship between indicator and program exists	5	5	
1.2	Indicators defined well and are data reported in an appropriate format	3	3	
1.3	The indicators accurately represent reality	3	3	
1.4	Indicators have a verifiable source	2	2	
1.5	Indicator presents all necessary data	2	2	
	TOTAL	15	15	1.00
2	Reliability			
2.1	Data collected and reported consistently over time, across project activities, and regions	5	5	
2.2	Data collection and reporting process guards against bias	5	4	
2.3	Process is reviewed periodically	4	4	
2.4	Random checks at each stage	4	3	
2.5	Data collection process is documented and problems are reported	2	2	
	TOTAL	20	18	0.90
3	Timeliness			
3.1	Data collected and reported on a frequent basis	5	5	
3.2	Reported data are sufficiently up to date	5	5	
3.3	Date of data collection is clearly identified	5	5	
	TOTAL	15	15	1.00
4	Precision			
4.1	Margin of error is reported	5	5	
4.2	Margin of error does not change the indicator	4	4	
4.3	Margin of error acceptable for decision-making	5	5	
4.4	Program targets include margin of error	4	4	
	TOTAL	18	18	1.00
5	Objectivity			
5.1	Data not subject to political and/or personal manipulation	5	4	
5.2	Independence of data collection, management, and assessment	5	5	
5.3	Impartial review of data collection process	4	4	
	TOTAL	14	13	0.93
6	Adequacy:			
6.1	Indicator fully measures intended outcome	5	4	
6.2	Indicator sufficiently characterizes outcome	5	5	
	TOTAL	10	9	0.90
7	Practicality			
7.1	Data collection process is viable	5	4	
7.2	Data collection not overly costly compared to the adequacy of the indicator	5	4	
7.3	Data collection is feasible	5	5	
	TOTAL	15	13	0.87
OVERALL RATING				6.60

Table A.8: Rating of Travel Time Indicator

	Criteria	Potential Score (1-5)	Actual Score (1-5)	Overall Rating
1	Validity			
1.1	Relationship between indicator and program exists	5	4	
1.2	Indicators defined well and are data reported in an appropriate format	3	2	
1.3	The indicators accurately represent reality	3	2	
1.4	Indicators have a verifiable source	2	2	
1.5	Indicator presents all necessary data	2	2	
	TOTAL	15	12	
2	Reliability			
2.1	Data collected and reported consistently over time, across project activities, and regions	5	3	
2.2	Data collection and reporting process guards against bias	5	4	
2.3	Process is reviewed periodically	4	3	
2.4	Random checks at each stage	4	3	
2.5	Data collection process is documented and problems are reported	2	2	
	TOTAL	20	15	
3	Timeliness			
3.1	Data collected and reported on a frequent basis	5	5	
3.2	Reported data are sufficiently up to date	5	5	
3.3	Date of data collection is clearly identified	5	5	
	TOTAL	15	15	
4	Precision			
4.1	Margin of error is reported	5	5	
4.2	Margin of error does not change the indicator	4	4	
4.3	Margin of error acceptable for decision-making	5	5	
4.4	Program targets include margin of error	4	4	
	TOTAL	18	18	
5	Objectivity			
5.1	Data not subject to political and/or personal manipulation	5	5	
5.2	Independence of data collection, management, and assessment	5	5	
5.3	Impartial review of data collection process	4	4	
	TOTAL	14	14	
6	Adequacy:			
6.1	Indicator fully measures intended outcome	5	5	
6.2	Indicator sufficiently characterizes outcome	5	5	
	TOTAL	10	10	
7	Practicality			
7.1	Data collection process is viable	5	5	
7.2	Data collection not overly costly compared to the adequacy of the indicator	5	5	
7.3	Data collection is feasible	5	5	
	TOTAL	15	15	
OVERALL RATING				6.55

Table A.9: Rating of Increase in Gross Revenue of Portfolio Companies (PCs) Indicator

	Criteria	Potential Score (1-5)	Actual Score (1-5)	Overall Rating
1	Validity			
1.1	Relationship between indicator and program exists	5	4	
1.2	Indicators defined well and are data reported in an appropriate format	3	3	
1.3	The indicators accurately represent reality	3	2	
1.4	Indicators have a verifiable source	2	2	
1.5	Indicator presents all necessary data	2	2	
	TOTAL	15	13	0.87
2	Reliability			
2.1	Data collected and reported consistently over time, across project activities, and regions	5	5	
2.2	Data collection and reporting process guards against bias	5	4	
2.3	Process is reviewed periodically	4	4	
2.4	Random checks at each stage	4	3	
2.5	Data collection process is documented and problems are reported	2	2	
	TOTAL	20	18	0.90
3	Timeliness			
3.1	Data collected and reported on a frequent basis	5	5	
3.2	Reported data are sufficiently up to date	5	5	
3.3	Date of data collection is clearly identified	5	5	
	TOTAL	15	15	1.00
4	Precision			
4.1	Margin of error is reported	5	5	
4.2	Margin of error does not change the indicator	4	4	
4.3	Margin of error acceptable for decision-making	5	5	
4.4	Program targets include margin of error	4	4	
	TOTAL	18	18	1.00
5	Objectivity			
5.1	Data not subject to political and/or personal manipulation	5	4	
5.2	Independence of data collection, management, and assessment	5	4	
5.3	Impartial review of data collection process	4	4	
	TOTAL	14	12	0.86
6	Adequacy:			
6.1	Indicator fully measures intended outcome	5	5	
6.2	Indicator sufficiently characterizes outcome	5	4	
	TOTAL	10	9	0.90
7	Practicality			
7.1	Data collection process is viable	5	5	
7.2	Data collection not overly costly compared to the adequacy of the indicator	5	5	
7.3	Data collection is feasible	5	5	
	TOTAL	15	15	1.00
OVERALL RATING				6.52

Table A.10: Rating of Increase in PC Employees Indicator

	Criteria	Potential Score (1-5)	Actual Score (1-5)	Overall Rating
1	Validity			
1.1	Relationship between indicator and program exists	5	4	
1.2	Indicators defined well and are data reported in an appropriate format	3	3	
1.3	The indicators accurately represent reality	3	3	
1.4	Indicators have a verifiable source	2	2	
1.5	Indicator presents all necessary data	2	2	
	TOTAL	15	14	0.93
2	Reliability			
2.1	Data collected and reported consistently over time, across project activities, and regions	5	4	
2.2	Data collection and reporting process guards against bias	5	4	
2.3	Process is reviewed periodically	4	3	
2.4	Random checks at each stage	4	3	
2.5	Data collection process is documented and problems are reported	2	2	
	TOTAL	20	16	0.80
3	Timeliness			
3.1	Data collected and reported on a frequent basis	5	5	
3.2	Reported data are sufficiently up to date	5	5	
3.3	Date of data collection is clearly identified	5	5	
	TOTAL	15	15	1.00
4	Precision			
4.1	Margin of error is reported	5	5	
4.2	Margin of error does not change the indicator	4	4	
4.3	Margin of error acceptable for decision-making	5	5	
4.4	Program targets include margin of error	4	4	
	TOTAL	18	18	1.00
5	Objectivity			
5.1	Data not subject to political and/or personal manipulation	5	5	
5.2	Independence of data collection, management, and assessment	5	4	
5.3	Impartial review of data collection process	4	4	
	TOTAL	14	13	0.93
6	Adequacy:			
6.1	Indicator fully measures intended outcome	5	5	
6.2	Indicator sufficiently characterizes outcome	5	4	
	TOTAL	10	9	0.90
7	Practicality			
7.1	Data collection process is viable	5	5	
7.2	Data collection not overly costly compared to the adequacy of the indicator	5	4	
7.3	Data collection is feasible	5	5	
	TOTAL	15	14	0.93
OVERALL RATING				6.50

Table A.11: Rating of Increase in Local Supplies to the PCs Indicator

	Criteria	Potential Score (1-5)	Actual Score (1-5)	Overall Rating
1	Validity			
1.1	Relationship between indicator and program exists	5	4	
1.2	Indicators defined well and are data reported in an appropriate format	3	2	
1.3	The indicators accurately represent reality	3	2	
1.4	Indicators have a verifiable source	2	2	
1.5	Indicator presents all necessary data	2	2	
	TOTAL	15	12	0.80
2	Reliability			
2.1	Data collected and reported consistently over time, across project activities, and regions	5	4	
2.2	Data collection and reporting process guards against bias	5	4	
2.3	Process is reviewed periodically	4	4	
2.4	Random checks at each stage	4	3	
2.5	Data collection process is documented and problems are reported	2	2	
	TOTAL	20	17	0.85
3	Timeliness			
3.1	Data collected and reported on a frequent basis	5	5	
3.2	Reported data are sufficiently up to date	5	5	
3.3	Date of data collection is clearly identified	5	5	
	TOTAL	15	15	1.00
4	Precision			
4.1	Margin of error is reported	5	5	
4.2	Margin of error does not change the indicator	4	4	
4.3	Margin of error acceptable for decision-making	5	5	
4.4	Program targets include margin of error	4	4	
	TOTAL	18	18	1.00
5	Objectivity			
5.1	Data not subject to political and/or personal manipulation	5	5	
5.2	Independence of data collection, management, and assessment	5	5	
5.3	Impartial review of data collection process	4	4	
	TOTAL	14	14	1.00
6	Adequacy:			
6.1	Indicator fully measures intended outcome	5	4	
6.2	Indicator sufficiently characterizes outcome	5	4	
	TOTAL	10	8	0.80
7	Practicality			
7.1	Data collection process is viable	5	5	
7.2	Data collection not overly costly compared to the adequacy of the indicator	5	4	
7.3	Data collection is feasible	5	4	
	TOTAL	15	13	0.87
OVERALL RATING				6.32

Table A.12: Rating of Increase in Wages Paid to the PC Employees Indicator

	Criteria	Potential Score (1-5)	Actual Score (1-5)	Overall Rating
1	Validity			
1.1	Relationship between indicator and program exists	5	5	
1.2	Indicators defined well and are data reported in an appropriate format	3	3	
1.3	The indicators accurately represent reality	3	2	
1.4	Indicators have a verifiable source	2	2	
1.5	Indicator presents all necessary data	2	2	
	TOTAL	15	14	0.93
2	Reliability			
2.1	Data collected and reported consistently over time, across project activities, and regions	5	4	
2.2	Data collection and reporting process guards against bias	5	5	
2.3	Process is reviewed periodically	4	4	
2.4	Random checks at each stage	4	4	
2.5	Data collection process is documented and problems are reported	2	2	
	TOTAL	20	19	0.95
3	Timeliness			
3.1	Data collected and reported on a frequent basis	5	5	
3.2	Reported data are sufficiently up to date	5	5	
3.3	Date of data collection is clearly identified	5	5	
	TOTAL	15	15	1.00
4	Precision			
4.1	Margin of error is reported	5	5	
4.2	Margin of error does not change the indicator	4	4	
4.3	Margin of error acceptable for decision-making	5	5	
4.4	Program targets include margin of error	4	4	
	TOTAL	18	18	1.00
5	Objectivity			
5.1	Data not subject to political and/or personal manipulation	5	4	
5.2	Independence of data collection, management, and assessment	5	4	
5.3	Impartial review of data collection process	4	3	
	TOTAL	14	11	0.79
6	Adequacy:			
6.1	Indicator fully measures intended outcome	5	5	
6.2	Indicator sufficiently characterizes outcome	5	4	
	TOTAL	10	9	0.90
7	Practicality			
7.1	Data collection process is viable	5	5	
7.2	Data collection not overly costly compared to the adequacy of the indicator	5	4	
7.3	Data collection is feasible	5	4	
	TOTAL	15	13	0.87
OVERALL RATING				6.44

**Table A.13: Rating of Increase in Locally Sources Goods and Services
Paid Purchased by PCs Indicator**

	Criteria	Potential Score (1-5)	Actual Score (1-5)	Overall Rating
1	Validity			
1.1	Relationship between indicator and program exists	5	5	
1.2	Indicators defined well and are data reported in an appropriate format	3	3	
1.3	The indicators accurately represent reality	3	2	
1.4	Indicators have a verifiable source	2	3	
1.5	Indicator presents all necessary data	2	3	
	TOTAL	15	15	1.00
2	Reliability			
2.1	Data collected and reported consistently over time, across project activities, and regions	5	4	
2.2	Data collection and reporting process guards against bias	5	4	
2.3	Process is reviewed periodically	4	4	
2.4	Random checks at each stage	4	3	
2.5	Data collection process is documented and problems are reported	2	2	
	TOTAL	20	18	0.90
3	Timeliness			
3.1	Data collected and reported on a frequent basis	5	5	
3.2	Reported data are sufficiently up to date	5	5	
3.3	Date of data collection is clearly identified	5	5	
	TOTAL	15	15	1.00
4	Precision			
4.1	Margin of error is reported	5	5	
4.2	Margin of error does not change the indicator	4	4	
4.3	Margin of error acceptable for decision-making	5	5	
4.4	Program targets include margin of error	4	4	
	TOTAL	18	18	1.00
5	Objectivity			
5.1	Data not subject to political and/or personal manipulation	5	5	
5.2	Independence of data collection, management, and assessment	5	4	
5.3	Impartial review of data collection process	4	3	
	TOTAL	14	12	0.86
6	Adequacy:			
6.1	Indicator fully measures intended outcome	5	5	
6.2	Indicator sufficiently characterizes outcome	5	5	
	TOTAL	10	10	1.00
7	Practicality			
7.1	Data collection process is viable	5	5	
7.2	Data collection not overly costly compared to the adequacy of the indicator	5	4	
7.3	Data collection is feasible	5	5	
	TOTAL	15	14	0.93
OVERALL RATING				6.69

Table A.14: Rating of Jobs Created Indicator

	Criteria	Potential Score (1-5)	Actual Score (1-5)	Overall Rating
1	Validity			
1.1	Relationship between indicator and program exists	5	5	
1.2	Indicators defined well and are data reported in an appropriate format	3	3	
1.3	The indicators accurately represent reality	3	3	
1.4	Indicators have a verifiable source	2	2	
1.5	Indicator presents all necessary data	2	2	
	TOTAL	15	15	
2	Reliability			
2.1	Data collected and reported consistently over time, across project activities, and regions	5	4	
2.2	Data collection and reporting process guards against bias	5	4	
2.3	Process is reviewed periodically	4	3	
2.4	Random checks at each stage	4	3	
2.5	Data collection process is documented and problems are reported	2	2	
	TOTAL	20	16	
3	Timeliness			
3.1	Data collected and reported on a frequent basis	5	5	
3.2	Reported data are sufficiently up to date	5	5	
3.3	Date of data collection is clearly identified	5	5	
	TOTAL	15	15	
4	Precision			
4.1	Margin of error is reported	5	5	
4.2	Margin of error does not change the indicator	4	4	
4.3	Margin of error acceptable for decision-making	5	5	
4.4	Program targets include margin of error	4	4	
	TOTAL	18	18	
5	Objectivity			
5.1	Data not subject to political and/or personal manipulation	5	4	
5.2	Independence of data collection, management, and assessment	5	4	
5.3	Impartial review of data collection process	4	4	
	TOTAL	14	12	
6	Adequacy:			
6.1	Indicator fully measures intended outcome	5	4	
6.2	Indicator sufficiently characterizes outcome	5	4	
	TOTAL	10	8	
7	Practicality			
7.1	Data collection process is viable	5	5	
7.2	Data collection not overly costly compared to the adequacy of the indicator	5	4	
7.3	Data collection is feasible	5	4	
	TOTAL	15	13	
OVERALL RATING				6.32

Table A.15: Rating of Household Net Income Indicator

	Criteria	Potential Score (1-5)	Actual Score (1-5)	Overall Rating
1	Validity			
1.1	Relationship between indicator and program exists	5	5	
1.2	Indicators defined well and are data reported in an appropriate format	3	3	
1.3	The indicators accurately represent reality	3	3	
1.4	Indicators have a verifiable source	2	2	
1.5	Indicator presents all necessary data	2	2	
	TOTAL	15	15	1.00
2	Reliability			
2.1	Data collected and reported consistently over time, across project activities, and regions	5	4	
2.2	Data collection and reporting process guards against bias	5	5	
2.3	Process is reviewed periodically	4	4	
2.4	Random checks at each stage	4	4	
2.5	Data collection process is documented and problems are reported	2	2	
	TOTAL	20	19	0.95
3	Timeliness			
3.1	Data collected and reported on a frequent basis	5	5	
3.2	Reported data are sufficiently up to date	5	5	
3.3	Date of data collection is clearly identified	5	5	
	TOTAL	15	15	1.00
4	Precision			
4.1	Margin of error is reported	5	5	
4.2	Margin of error does not change the indicator	4	4	
4.3	Margin of error acceptable for decision-making	5	5	
4.4	Program targets include margin of error	4	4	
	TOTAL	18	18	1.00
5	Objectivity			
5.1	Data not subject to political and/or personal manipulation	5	5	
5.2	Independence of data collection, management, and assessment	5	5	
5.3	Impartial review of data collection process	4	4	
	TOTAL	14	14	1.00
6	Adequacy:			
6.1	Indicator fully measures intended outcome	5	5	
6.2	Indicator sufficiently characterizes outcome	5	5	
	TOTAL	10	10	1.00
7	Practicality			
7.1	Data collection process is viable	5	5	
7.2	Data collection not overly costly compared to the adequacy of the indicator	5	5	
7.3	Data collection is feasible	5	5	
	TOTAL	15	15	1.00
OVERALL RATING				6.95

Table A.16: Rating of Household Firm Income Indicator

	Criteria	Potential Score (1-5)	Actual Score (1-5)	Overall Rating
1	Validity			
1.1	Relationship between indicator and program exists	5	5	
1.2	Indicators defined well and are data reported in an appropriate format	3	3	
1.3	The indicators accurately represent reality	3	2	
1.4	Indicators have a verifiable source	2	2	
1.5	Indicator presents all necessary data	2	1	
	TOTAL	15	13	0.87
2	Reliability			
2.1	Data collected and reported consistently over time, across project activities, and regions	5	5	
2.2	Data collection and reporting process guards against bias	5	5	
2.3	Process is reviewed periodically	4	4	
2.4	Random checks at each stage	4	3	
2.5	Data collection process is documented and problems are reported	2	2	
	TOTAL	20	19	0.95
3	Timeliness			
3.1	Data collected and reported on a frequent basis	5	5	
3.2	Reported data are sufficiently up to date	5	5	
3.3	Date of data collection is clearly identified	5	5	
	TOTAL	15	15	1.00
4	Precision			
4.1	Margin of error is reported	5	5	
4.2	Margin of error does not change the indicator	4	4	
4.3	Margin of error acceptable for decision-making	5	5	
4.4	Program targets include margin of error	4	4	
	TOTAL	18	18	1.00
5	Objectivity			
5.1	Data not subject to political and/or personal manipulation	5	5	
5.2	Independence of data collection, management, and assessment	5	5	
5.3	Impartial review of data collection process	4	3	
	TOTAL	14	13	0.93
6	Adequacy:			
6.1	Indicator fully measures intended outcome	5	4	
6.2	Indicator sufficiently characterizes outcome	5	4	
	TOTAL	10	8	0.80
7	Practicality			
7.1	Data collection process is viable	5	5	
7.2	Data collection not overly costly compared to the adequacy of the indicator	5	5	
7.3	Data collection is feasible	5	5	
	TOTAL	15	15	1.00
OVERALL RATING				6.55

Table A.17: Rating of Beneficiaries Indicator

	Criteria	Potential Score (1-5)	Actual Score (1-5)	Overall Rating
1	Validity			
1.1	Relationship between indicator and program exists	5	5	
1.2	Indicators defined well and are data reported in an appropriate format	3	3	
1.3	The indicators accurately represent reality	3	3	
1.4	Indicators have a verifiable source	2	2	
1.5	Indicator presents all necessary data	2	2	
	TOTAL	15	15	1.00
2	Reliability			
2.1	Data collected and reported consistently over time, across project activities, and regions	5	5	
2.2	Data collection and reporting process guards against bias	5	5	
2.3	Process is reviewed periodically	4	4	
2.4	Random checks at each stage	4	4	
2.5	Data collection process is documented and problems are reported	2	2	
	TOTAL	20	20	1.00
3	Timeliness			
3.1	Data collected and reported on a frequent basis	5	5	
3.2	Reported data are sufficiently up to date	5	5	
3.3	Date of data collection is clearly identified	5	5	
	TOTAL	15	15	1.00
4	Precision			
4.1	Margin of error is reported	5	5	
4.2	Margin of error does not change the indicator	4	4	
4.3	Margin of error acceptable for decision-making	5	5	
4.4	Program targets include margin of error	4	4	
	TOTAL	18	18	1.00
5	Objectivity			
5.1	Data not subject to political and/or personal manipulation	5	5	
5.2	Independence of data collection, management, and assessment	5	5	
5.3	Impartial review of data collection process	4	4	
	TOTAL	14	14	1.00
6	Adequacy:			
6.1	Indicator fully measures intended outcome	5	5	
6.2	Indicator sufficiently characterizes outcome	5	5	
	TOTAL	10	10	1.00
7	Practicality			
7.1	Data collection process is viable	5	5	
7.2	Data collection not overly costly compared to the adequacy of the indicator	5	5	
7.3	Data collection is feasible	5	5	
	TOTAL	15	15	1.00
OVERALL RATING			7.00	

APPENDIX B: RATING TABLES FOR SURVEYS

Table B1: ADA Direct Beneficiary Survey

	Criteria	Potential Score (1-5)	Actual Score (1-5)	Overall Rating
1	Validity			
1.1	Relationship between variables and program exists	5	5	
1.2	Variables defined well and are data reported in an appropriate format	4	4	
1.3	The variables accurately represent reality	5	5	
1.4	Variables have a verifiable source	2	2	
1.5	Variables present all necessary data	2	2	
	TOTAL	18	18	
2	Reliability			
2.1	Data collected and reported consistently over time, across project activities, and regions	5	5	
2.2	Data collection and reporting process guards against bias	5	5	
2.3	Process is reviewed periodically	4	4	
2.4	Random checks at each stage	3	3	
2.5	Data collection process is documented and problems are reported	3	3	
	TOTAL	20	20	
3	Timeliness			
3.1	Data collected and reported on a frequent basis	5	4	
3.2	Reported data are sufficiently up to date	5	5	
3.3	Date of data collection is clearly identified	5	5	
	TOTAL	15	14	
4	Precision			
4.1	Adequate data collected to report margins of error	5	5	
4.2	Margins of error do not change the variables	5	4	
4.3	Margins of error acceptable for program indicator construction	5	5	
4.4	Margins of error do not change program indicators	5	5	
	TOTAL	20	19	
5	Objectivity			
5.1	Data are not subject to political and/or personal manipulation	5	5	
5.2	Independence of data collection, management, and assessment	5	5	
5.3	Impartial review of data collection process	5	4	
	TOTAL	15	14	.93
6	Adequacy:			
6.1	Variables fully measure intended characteristics and outcome	5	5	
6.2	Variables sufficiently describes characteristics and outcomes	5	5	
	TOTAL	10	10	1.00
7	Practicality			
7.1	Data collection process is viable	5	5	
7.2	Data collection not overly costly compared to the adequacy of the variables	5	5	
7.3	Data collection is feasible	5	5	
	TOTAL	15	15	1.00
OVERALL RATING				6.81

Table B2: Settlement Infrastructure Survey (SIS)

	Criteria	Potential Score (1-5)	Actual Score (1-5)	Overall Rating
1	Validity			
1.1	Relationship between variables and program exists	5	5	
1.2	Variables defined well and are data reported in an appropriate format	4	4	
1.3	The variables accurately represent reality	5	5	
1.4	Variables have a verifiable source	2	2	
1.5	Variables present all necessary data	2	2	
	TOTAL	18	18	1.00
2	Reliability			
2.1	Data collected and reported consistently over time, across project activities, and regions	5	5	
2.2	Data collection and reporting process guards against bias	5	5	
2.3	Process is reviewed periodically	4	4	
2.4	Random checks at each stage	3	3	
2.5	Data collection process is documented and problems are reported	3	3	
	TOTAL	20	20	1.00
3	Timeliness			
3.1	Data collected and reported on a frequent basis	5	3	
3.2	Reported data are sufficiently up to date	5	4	
3.3	Date of data collection is clearly identified	5	5	
	TOTAL	15	12	0.80
4	Precision			
4.1	Adequate data collected to report margins of error	5	5	
4.2	Margins of error do not change the variables	5	5	
4.3	Margins of error acceptable for program indicator construction	5	5	
4.4	Margins of error do not change program indicators	5	5	
	TOTAL	20	20	1.00
5	Objectivity			
5.1	Data are not subject to political and/or personal manipulation	5	5	
5.2	Independence of data collection, management, and assessment	5	5	
5.3	Impartial review of data collection process	5	5	
	TOTAL	15	15	1.00
6	Adequacy:			
6.1	Variables fully measure intended characteristics and outcome	5	5	
6.2	Variables sufficiently describes characteristics and outcomes	5	5	
	TOTAL	10	10	1.00
7	Practicality			
7.1	Data collection process is viable	5	4	
7.2	Data collection not overly costly compared to the adequacy of the variables	5	5	
7.3	Data collection is feasible	5	5	
	TOTAL	15	14	0.93
OVERALL RATING				6.73

Table B3: Road Users Survey

	Criteria	Potential Score (1-5)	Actual Score (1-5)	Overall Rating
1	Validity			
1.1	Relationship between variables and program exists	5	5	
1.2	Variables defined well and are data reported in an appropriate format	4	4	
1.3	The variables accurately represent reality	5	5	
1.4	Variables have a verifiable source	2	2	
1.5	Variables present all necessary data	2	2	
	TOTAL	18	18	1.00
2	Reliability			
2.1	Data collected and reported consistently over time, across project activities, and regions	5	5	
2.2	Data collection and reporting process guards against bias	5	5	
2.3	Process is reviewed periodically	4	4	
2.4	Random checks at each stage	3	3	
2.5	Data collection process is documented and problems are reported	3	3	
	TOTAL	20	20	1.00
3	Timeliness			
3.1	Data collected and reported on a frequent basis	5	4	
3.2	Reported data are sufficiently up to date	5	5	
3.3	Date of data collection is clearly identified	5	5	
	TOTAL	15	14	0.93
4	Precision			
4.1	Adequate data collected to report margins of error	5	4	
4.2	Margins of error do not change the variables	5	5	
4.3	Margins of error acceptable for program indicator construction	5	5	
4.4	Margins of error do not change program indicators	5	5	
	TOTAL	20	19	0.95
5	Objectivity			
5.1	Data are not subject to political and/or personal manipulation	5	4	
5.2	Independence of data collection, management, and assessment	5	5	
5.3	Impartial review of data collection process	5	5	
	TOTAL	15	14	0.93
6	Adequacy:			
6.1	Variables fully measure intended characteristics and outcome	5	5	
6.2	Variables sufficiently describes characteristics and outcomes	5	5	
	TOTAL	10	10	1.00
7	Practicality			
7.1	Data collection process is viable	5	4	
7.2	Data collection not overly costly compared to the adequacy of the variables	5	5	
7.3	Data collection is feasible	5	5	
	TOTAL	15	14	0.93
OVERALL RATING				6.74

Table B4: Village Infrastructure Census (VIC)

	Criteria	Potential Score (1-5)	Actual Score (1-5)	Overall Rating
1	Validity			
1.1	Relationship between variables and program exists	5	5	
1.2	Variables defined well and are data reported in an appropriate format	4	4	
1.3	The variables accurately represent reality	5	5	
1.4	Variables have a verifiable source	2	2	
1.5	Variables present all necessary data	2	2	
	TOTAL	18	18	1.00
2	Reliability			
2.1	Data collected and reported consistently over time, across project activities, and regions	5	5	
2.2	Data collection and reporting process guards against bias	5	5	
2.3	Process is reviewed periodically	4	4	
2.4	Random checks at each stage	3	3	
2.5	Data collection process is documented and problems are reported	3	3	
	TOTAL	20	20	1.00
3	Timeliness			
3.1	Data collected and reported on a frequent basis	5	5	
3.2	Reported data are sufficiently up to date	5	5	
3.3	Date of data collection is clearly identified	5	5	
	TOTAL	15	15	1.00
4	Precision			
4.1	Adequate data collected to report margins of error	5	5	
4.2	Margins of error do not change the variables	5	5	
4.3	Margins of error acceptable for program indicator construction	5	5	
4.4	Margins of error do not change program indicators	5	5	
	TOTAL	20	20	1.00
5	Objectivity			
5.1	Data are not subject to political and/or personal manipulation	5	5	
5.2	Independence of data collection, management, and assessment	5	5	
5.3	Impartial review of data collection process	5	5	
	TOTAL	15	15	1.00
6	Adequacy:			
6.1	Variables fully measure intended characteristics and outcome	5	5	
6.2	Variables sufficiently describes characteristics and outcomes	5	5	
	TOTAL	10	10	1.00
7	Practicality			
7.1	Data collection process is viable	5	4	
7.2	Data collection not overly costly compared to the adequacy of the variables	5	5	
7.3	Data collection is feasible	5	5	
	TOTAL	15	14	0.93
OVERALL RATING				6.93

Table B5: Integrated Household Survey (IHS)

	Criteria	Potential Score (1-5)	Actual Score (1-5)	Overall Rating
1	Validity			
1.1	Relationship between variables and program exists	5	5	
1.2	Variables defined well and are data reported in an appropriate format	4	4	
1.3	The variables accurately represent reality	5	5	
1.4	Variables have a verifiable source	2	2	
1.5	Variables present all necessary data	2	2	
	TOTAL	18	18	1.00
2	Reliability			
2.1	Data collected and reported consistently over time, across project activities, and regions	5	5	
2.2	Data collection and reporting process guards against bias	5	5	
2.3	Process is reviewed periodically	4	4	
2.4	Random checks at each stage	3	3	
2.5	Data collection process is documented and problems are reported	3	3	
	TOTAL	20	20	1.00
3	Timeliness			
3.1	Data collected and reported on a frequent basis	5	5	
3.2	Reported data are sufficiently up to date	5	5	
3.3	Date of data collection is clearly identified	5	5	
	TOTAL	15	15	1.00
4	Precision			
4.1	Adequate data collected to report margins of error	5	5	
4.2	Margins of error do not change the variables	5	5	
4.3	Margins of error acceptable for program indicator construction	5	5	
4.4	Margins of error do not change program indicators	5	5	
	TOTAL	20	20	1.00
5	Objectivity			
5.1	Data are not subject to political and/or personal manipulation	5	4	
5.2	Independence of data collection, management, and assessment	5	5	
5.3	Impartial review of data collection process	5	5	
	TOTAL	15	14	0.93
6	Adequacy:			
6.1	Variables fully measure intended characteristics and outcome	5	5	
6.2	Variables sufficiently describes characteristics and outcomes	5	5	
	TOTAL	10	10	1.00
7	Practicality			
7.1	Data collection process is viable	5	5	
7.2	Data collection not overly costly compared to the adequacy of the variables	5	5	
7.3	Data collection is feasible	5	5	
	TOTAL	15	15	1.00
OVERALL RATING				6.93