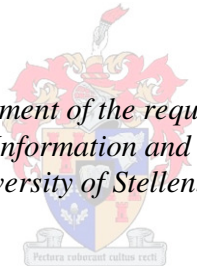
The Relation between ICT and Poverty Reduction: The Central Statistical Agency of Ethiopia

by
Ermyas Arega Tessema



Thesis presented in fulfilment of the requirements for the degree of Master of Philosophy (Information and Knowledge Management University of Stellenbosch

Supervisor: Daniel F. Botha
Faculty of Arts and Social Sciences
Department of Information Science

December 2010

DECLARATION

By submitting this thesis/dissertation electronically, I declare that the entirety of the work contained therein is my own, original work, and that I have not previously, in its entirety or in part, submitted it for obtaining any qualification.

December 2010

Copyright © 2010 University of Stellenbosch

All rights reserved

ABSTRACT

National Statistical offices (NSOs) are the sources of wide ranges of socio-economic, demographic and agricultural data and information that are used to monitor and evaluate development programs and formulate policies. The data generated by NSOs is used as basis for making decisions and also used to assess the extent and causes of poverty. Various stakeholders such as researchers, the World Bank, the International Monetary Fund, the UN, and various NGOs prepare and release research materials and annual reports using data and information obtained from NSOs and line ministries. For example, Deneulin and Shahani state that one of the intentions of the annual Human Development Report (HDR) prepared by the UNDP is "*to assess the quality of life of a population and be an advocacy tool for its improvement with a political purpose of raising awareness and generating debate on public issues and concerns which would otherwise not be on the political agenda*"[1]. Based on the different approaches to poverty, different sets of data and information are produced and used for poverty measurement. Mostly, poverty is measured using data obtained from nationally representative household surveys which focus on income and expenditure, ownership, access to and use of some basic services. Another approach uses data on mental satisfaction; still others assume poverty to be multi-dimensional and argue that income alone is not enough. They view poverty as deprivation of basic capabilities due to high rates of mortality, illiteracy, malnourishment, unemployment, ill health, lack of education and social exclusion, etc[2]. The quality of data and information (such as integrity, methodological soundness, accuracy and reliability, serviceability and accessibility) generated by data-producing-agencies therefore needs to be preserved and improved in order to obtain meaningful results from the measurement of poverty in any of the approaches and to satisfy the growing data quality demands of stakeholders. Loshin states that "*strategic decisions based on untrustworthy information are likely to result in poor decisions*"[3]. This study focuses on the role played by national statistical offices in poverty reduction in general. It examines the various activities, players, interactions, and ICTs used at the various stages of the statistical

[1] Alkire and Shani 2009, p.24
[2] Sen 1999, pp. 103,109
[3] Loshin 2001, p.10

process in the Ethiopian Central Statistical Agency (CSA) to generate poverty-related data and information and how the quality of this data can be preserved and improved.

The purpose of this research is therefore to identify poverty related data quality problems with respect to the IMF's DQAF and assess where in the statistical process specific types of ICTs can improve data quality. For this reason interpretive case study method with the researcher as participant observer was adopted to study how poverty related data and information is produced. It was found out that some of the data quality problems can be addressed using appropriate ICTs with the availability of reliable power infrastructures.

OPSOMMING

Nasionale Statistiekkantore (NSOs) is die bron van ‘n wye reeks sosio-ekonomiese, demografiese en landboukundige data en inligting wat gebruik word om ontwikkelingsprogramme te monitor en te evalueer. Die data wat deur NSOs geskep word, word aangewend as grondslag vir besluitneming. Die data word ook gebruik om die omvang en oorsake van armoede te bepaal. Verskeie betrokkenes soos navorsers, die Wêreldbank (WB), Internasionale Monetêre Fonds (IMF) en die VN en NSOs skep en versprei verskillende navorsingsuitsette en jaarverslae wat gebruik maak van die data en inligting wat verkry word van NSOs en ministeries. So konstateer Deneulin en Shahani dat een van die doelstellings van die Verslag op Menslike Ontwikkeling (HDR), soos opgestel deur die VNDP, is om “*die lewensgehalte van ‘n bevolking te skat en om as werktuig en voorspraak vir die verbetering daarvan op te tree, met die politiese doelwit om bewustheid te verhoog en debatvoering oor openbare sake en kwessies, wat andersins nie op die agenda sou verskyn nie, aan te voor”*.[4]. Na gelang van die verskillende benaderings tot armoede word verskillende stelle data en inligting geproduseer en gebruik vir die meting van armoede. Armoede word gewoonlik gemeet deur data te gebruik wat bekom word van landswye opnames van huishoudings en ingestel is op inkomste en besteding, besitreg, toegang tot en die gebruik van ‘n paar basiese dienste. ‘n Ander benadering gebruik data oor geestelike bevrediging; ander weer aanvaar dat armoede multidimensioneel is en voer aan dat inkomste alleen nie genoeg is nie. Hulle beskou armoede as die ontbering van basiese vermoëns weens ‘n hoë sterftesyfer, ongeletterdheid, ondervoeding, siekte, gebrekkige onderwys, sosiale uitsluiting en dies meer[5]. Die gehalte van data en inligting (soos integriteit, metodologiese deeglikheid, akkuraatheid en betroubaarheid, bruikbaarheid en toeganklikheid) wat deur agentskappe opgelewer word moet dus bewaar en verbeter word ten einde ‘n beduidende resultaat te bekom uit die meting van armoede deur enige van die benaderings en ook om belanghebbendes se groeiende aandrang op datagehalte te bevredig. Loshin beweer dat “*strategiese besluite gebaseer op onbetroubare inligting waarskynlik swak besluitneming tot gevolg sal hê”*.[6]. Hierdie ondersoek konsentreer op die rol wat gespeel word deur nasionale statistiekkantore in die algemene bekamping van armoede. Dit ondersoek die verskillende aktiwiteite, rolspelers, interaksies en ICTs wat op verskeie stadiums van die statistiese proses deur die Etiopiese Sentrale Statistiekagentskap (CSA) gebruik word om data en inligting oor armoede te skep en hoe die betroubaarheid van data behou en verbeter kan word.

[4] Alkire en Shani 2009, bl. 24
[5] Sen 1999, bl. 103,109
[6] Loshin 2001, bl.10

Die doel van hierdie navorsing is dus om kwaliteitsprobleme wat verband hou met armoededata ten opsigte van die IMF se DQAF te identifiseer en om te besluit waar in die statistiese proses bepaalde soort ICT's die gehalte van data kan identifiseer. Om hierdie rede is die interpretiewe gevallestudie-metode aanvaar om te bepaal hoe armoede-verwante data en inligting geskep word. Die slotsom was dat sommige van die probleme in datagehalte aangespreek kan word deur die gebruik van gepaste ICT's met die beskikbaarheid van betroubare mag-infrakstrukture.

DEDICATION

To my *parents*, *brothers* and *sisters* - I don't have enough words to thank them for their unlimited support and encouragement**.**

ACKNOWLEDGEMENTS

I would like to extend my sincere gratitude to my supervisor Mr D F Botha for his insightful advice and guidance on the research. I was really inspired and motivated during the course of the research. I also would like to mention Dr. Hans Müller, for without his continuous support during my studies at Stellenbosch, the research would not have been completed. I deeply appreciate his guidance.

TABLE OF CONTENTS

DECLARATION ii

ABSTRACT iii

DEDICATION vii

ACKNOWLEDGEMENTS viii

TABLE OF CONTENTS ix

ABBREVIATIONS xii

LIST OF FIGURES AND TABLES xiii

INTRODUCTION 1

1.1 Introduction 1

1.2 Background 1

1.3 Research Problem 4

1.4 Research Question 5

1.5 Research Objective 6

1.6 Structure of the Study 8

1.7 Conclusion 8

RESEARCH METHODOLOGY 9

2.1 Introduction 9

2.2 Research Method 9

2.3 Major References 15

2.4 Rationale for the Research Topic 17

2.5 Conclusion 18

POVERTY AND POVERTY REDUCTION 20

3.1 Introduction 20

3.2 Overview of Poverty 20

3.3 Trends of Poverty 28
3.4 Poverty Indicators 37
3.4.1 What are Indicators? 37
3.4.2 MDGs Indicators 40
3.4.3 Human Development and Human Poverty Indicator 42
3.4.4 ICTs Indicators 47
3.4.5 WMS and HICES Indicators 48
3.5 Poverty Reduction Strategy Papers (PRSPs) 49
3.6 Conclusion 51
DATA QUALITY AND ICT 54
4.1 Introduction 54
4.2 Data Quality 54
4.3 Data, Information and Knowledge 60
4.4 ICTs in Improving Data Quality 63
4.5 Conclusion 81
CENTRAL STATISTICAL AGENCY OF ETHIOPIA AND ITS ACTIVITIES 83
5.1 Introduction 83
5.2 Role of CSA and Stakeholders 84
5.3 Role of the CSA in PRSP 88
5.4 ICT Strategies of the Government and CSA 90
5.4.1 ICT 90
5.4.2 Telecommunications 90
5.4.3 Strategies of the CSA 93
5.5 ICT Infrastructure of CSA 94
5.6 Conclusion 102
ANALYSIS 103
6.1 Introduction 103
6.2 Flow Charts 103

6.3 Description of Activities in Generating Data and Information 110
6.4 Conclusion 121
CONCLUSION AND RECOMMENDATION 122
7.1 Conclusion 122
7.2 Recommendation 125
7.2.1 Mobile Data Collection 125
7.2.2 Potential of Using Existing ICT Infrastructure to Other Purposes 132
BIBLIOGRAPHY 134
APPENDIX 141
1. Appendix A 141
2. Appendix B 146
3. Appendix C 149

ABBREVIATIONS

CSA	Central Statistical Agency (of Ethiopia)
CAPI	Computer Assisted Personal Interviewing
DP	Data Processing
EA	Enumeration Area
EICTDA	Ethiopian ICT Development Agency
HDI	Human Development Index
HDR	Human Development Report
HICES	Household Income and Expenditure Survey
GIS	Geographic Information System
ICT	Information and Communications Technologies
ICT4D	Information and Communication Technologies for Development (of Ethiopia)
IMF	International Monetary Fund
ITU	International Telecommunications Union
KM	Knowledge Management
MDG	Millennium Development Goals
MoCB	Ministry of Capacity Development (of Ethiopia)
MoFED	Ministry of Finance and Economic Development (of Ethiopia)
NGO	Non-Governmental Organisations
NSO	National Statistical Offices
PAPI	Paper and Pencil Interview
PDA	Personal Digital Assistant
PRSP	Poverty Reduction Strategy Paper
PASDEP	Plan for Accelerated and Sustained Development to End Poverty
SDPRP	Sustainable Development and Poverty Reduction Paper (of Ethiopia)
SMS	Subject matter specialists
UMPC	Ultra Mobile Pocket PC
UNDP	United Nations Development Program
UNECA	United Nations Economic Commission for Africa
USAID	United States Agency for International Development
WB	World Bank
WDI	World Development Indicators
WMS	Welfare Monitoring Survey

LIST OF FIGURES AND TABLES

Figures

Figure 3-1 Intermediate and final indicators....40
Figure 4-1 Components of an information system 64
Figure 6-1 General activities at Head Office and the branch offices.....105
Figure 6-2 Detailed questionnaire and data flow in the branch offices 106
Figure 6-3 General questionnaire and data flow in the data processing department at the head office. 107
Figure 6-4 Flow chart of activities during data entry 108
Figure 6-5 Flow chart of activities during data cleaning 109
Figure 7-1 Data Flow under mobile data collection 130

Tables

Table 1-1 Summary of research methodology....15
Table 3-1 Estimates of people who live under poverty at regional levels 30
Table 3-2 Trends in total poverty indices at national level and by rural and urban areas. 33
Table 3-3 Trends in food poverty indices at national level and by rural and urban areas 34
Table 4-1 Data quality problems by publication, data source, and informational base....59
Table 4-2 Mean time and median distance travelled per day by interviewers for mapping and interview 72
Table 4-3 Software and hardware problems encountered during CAPI mode of data collection using PDAs 74
Table 4-4 Effects of CAPI and PAPI modes on data quality....77
Table 4-5 IS components and data quality in the different stages of the statistical process....81
Table 5-1 The ICT development indicator for Ethiopia 92
Table 5-2 Major ICTs and their usage in CSA 97
Table 7-1 Possible Data Quality Improvements with Mobile Data collection 129

CHAPTER ONE

INTRODUCTION

1.1 Introduction

In this chapter the research problem, issues arising from it and overall objectives are discussed. They are put in the context of poverty reduction, linking data producing agencies with other stakeholders involved in poverty analysis, preparation of poverty reduction strategy papers (PRSPs), anti-poverty programs, and decision-making. Stakeholders in this study are generally seen as internal (which includes the employees at all levels in a data-producing agency such as a national statistical office, who are actually involved in the generation of data and information), and external (which includes various line ministries, international institutions, non-governmental organizations or NGOs, professional associations, researchers, respondents, etc). The quality of data and information is seen as important and the main thrust of the research is therefore to study the data and information generation process in order to apply information and communications technologies (ICTs) to preserve and improve the quality of data and information at the various stages of the statistical process.

1.2 Background

Poverty has many faces – the way it manifests itself in different countries and nations is correspondingly different. In the developing world poverty may manifest itself as the inability of individuals or households to exceed a minimum threshold of food intake, or to gain access to basic needs and services such as potable water, health and communication (roads and telephone), whereas in the developed world poverty would be defined by a completely different set of criteria.

In order to determine levels of poverty in a country or region, concerned government institutions – mainly the national statistical offices – conduct both quantitative and qualitative

surveys to determine the numbers or percentages of individuals who exist below or above a certain poverty threshold. Other stakeholders use these surveys and combine the information in them with other studies obtained from various other line ministries, research institutions and NGOs to formulate evidence-based policy decisions, devise intervention programs and monitor and evaluate such programs. Evidence-based policy is defined as "*an approach that helps people make well-informed decisions about policies, programmes and projects by putting the best available evidence from research at the heart of policy development and implementation."*[7] Identifying the poor in order to develop and implement anti-poverty strategies requires continuous monitoring and evaluation. Different stakeholders such as national statistical offices (NSOs), sectoral Government ministries, local and international NGOs, international organizations such as the World Bank (WB) and the International Monetary Fund (IMF), donors, etc participate in this under national and international initiatives. One of the international initiatives often cited is the Millennium Development Goals (MDGs).

According to the United Nations (UN)[8], the MDGs are the development aspirations of the international community, which have been adopted by over 190 countries as "*not only development objectives; they encompass universally accepted human values and rights such as freedom from hunger, the right to basic education, the right to health and a responsibility to future generations*."

Goal Six of the MDGs for example is combating diseases such as HIV/AIDS and malaria. In this report it was found that there was positive progress towards this goal thanks to successful intervention programs such as improvements in prevention programs, expansion of HIV/Aids treatment services and the delivery of bed-nets, drugs and vaccines[9]. Such interventions that identify and target vulnerable or affected people require coordinated activities from all stakeholders at all levels in policy formulation, implementation, monitoring and evaluation. National Statistical Offices, as one of the stakeholders, conduct nationally representative poverty-related household surveys and generate data and information that enable other

[7] Nutley et al, 2008, p.13
[8] MDG Report, 2008
[9] MDG Report, 2008, p.28

stakeholders to monitor and evaluate existing programs and to identify the extent and causes of poverty.

Other initiatives include ICT for Development (ICT4D) both at the global and national levels. Different countries, United Nations agencies (UNDP, UNECA), and development partners such as the European Union (EU), Departments for International Development (DFIDs), the US Agency for International Development (USAID). International organizations such as the WB and IMF have their own ICT4D initiatives and programmes. The ICT4D initiative of the Ethiopian Government is aimed at bringing about all-inclusive economic development by using ICTs as enablers in different sectors. According to the national ICT4D five-year action plan[10], the aim of the national ICT policy is to build an information- and knowledge-based society and harness the benefits of ICT to transform the economy from an agriculture-based to a knowledge-based economy. The Central Statistical Agency (CSA) of Ethiopia, as a federal agency collects, processes, stores, and disseminates nationally representative socio-economic, demographic, and agricultural data and information, which is used by stakeholders such as researchers, students and other government and non-government organizations. These data and information are central to the programs of the ICT4D, whose ultimate goal is implementing

> *government-to-citizens (G2C), government-to-business (G2B) and government-to-government (G2G) e-government programmes and initiatives that will facilitate improvements in information provision to citizens and businesses, service delivery to citizens and businesses and improvements in government administration*[11].

Most of the data used by international organizations, donors, and other stakeholders who follow up the progress made by countries towards meeting the MDGs, and other initiatives such as access to and use of ICTs, is collected by national statistical offices and other line ministries. National statistical offices also collect administrative record information from line ministries and other government organizations, which makes them one-stop centres for the country's data and information. The ideal is to ensure that the quality of this data is retained and enhanced in such a way that it is accurate and reliable, timely, relevant, consistent, clear

[10]MoCB, 2006, p.8.
Can also be accessed from the Ethiopian ICT Development Agency (EICTDA). *The National ICT for development five years action plan for Ethiopia (2006-2010).* Retrieved June 1, 2009 from http://www.eictda.gov.et/?q=node/15
[11] MoCB, 2006, p.26

and understandable, accessible, and that assistance from data-producing agencies regarding the nature and use of the data is always available to stakeholders.

1.3 Research Problem

Poverty reduction at the global, regional and country level is a task that involves the concerted and coordinated effort of different stakeholders. Data and information obtained from data producing agencies such as National Statistical Offices from household and establishment surveys and other line ministries in the form of administrative records, are valuable inputs for stakeholders engaged in poverty analysis and needing to make decisions based on the best available (and most reliable) information in the development of PRSPs and anti-poverty programs, in monitoring and evaluating these programs to see how far countries have advanced in meeting the goals set in international initiatives such as the MDGs. The NSOs of different countries play a vital role in this coordinated effort in fighting poverty by providing the necessary data and information.

Sometimes external stakeholders to NSOs such as the WB or UN do not include the progress or lack thereof made by some countries in combating poverty in their annual reports due to the state of macro- and micro-data. Understanding the nature of the problems related to data quality enables them to provide solutions so that external stakeholders are able to make more informed decisions and to develop effective anti-poverty programs. Data quality, according to IMF's data quality assessment framework (DQAF), includes "*integrity*, *methodological soundness*, *accuracy and reliability*, *serviceability*, and *accessibility"*. In NSOs, an understanding of the underlying processes and interactions and the flow and production of data and information, enables them to understand the problems encountered at the various steps of the statistical process and allows them to provide possible solutions to these problems through the application of ICTs where and, where necessary, according to stakeholders' data quality demands. The rationale for selecting this particular research subject is further described in Section Three of Chapter 2. In order to understand and provide solutions to this problem, the research question and sub-questions are outlined in the next section.

1.4 Research Question

Primary question:

Is there a significant relationship between the use of ICT in the Central Statistical Agency of Ethiopia (CSA) and its ability to reduce poverty?

Governments and international institutions use data generated by national statistical offices and other government line ministries in the form of periodic surveys such as welfare and monitoring surveys (WMS), and household income consumption and expenditure surveys (HICES) to monitor and evaluate the effectiveness of anti-poverty programs, to estimate the national proportion of the poor, and to monitor trends in poverty, among other things. This study describes the problems relating to poverty-related data quality from household and establishment surveys conducted by statistical offices from the point of view of both internal and external stakeholders. It investigates how data and information are generated by studying the use of ICTs, the role players involved and their interactions, the data-generation process, and the flow of data and information in the various stages of the statistical processes at the CSA of Ethiopia.

It also tries to establish the relationship between the various ICTs that are used during pre-data collection, during data collection, and during post-data collection stages of the statistical processes and data quality.

Some subsidiary questions need to be answered in order to answer the primary question:

- How are data and information (the informational basis for poverty) generated by national statistical offices (e.g. by the CSA of Ethiopia)?
- What are the different activities performed and how do different internal role players (the employees of CSA) interact to produce data and information?
- What are the different ICTs used in the process of producing data and information (during the pre-data collection, data collection, and post-data collection stages) by national statistical offices in general and the CSA in particular?
- How can the quality of the data be maintained and improved at various stages of the statistical process?

- How can barriers (such as inadequate computer literacy and competence, telecommunication and electricity infrastructures) affect the use of ICTs (i.e. mainly in the form of field data collection by means of portable PCs) and how can they be addressed?

Unit of Analysis:

- The process or flow of data and information,
- The actors involved and activities performed, and
- The resources used

The CSA of Ethiopia is an agency mandated to collect, process, and disseminate data and information. This study therefore focuses on data and information flow, role players and activities, and resources that are implemented to produce data and information at various stages of the statistical process in the CSA.

1.5 Research Objective

Preserving and improving the quality of the data and information generated by data producing agencies such as NSOs and other government line ministries is vital for external stakeholders such as economic planners and international financial institutions who are involved in assessing the extent and causes of poverty and in developing poverty-related strategies and programs.

The CSA of Ethiopia, as stated in previous sections collects, processes, stores, and disseminates nationally representative socio-economic, demographic, and agricultural data and information. It also compiles administrative records from line ministries and other government agencies. Various other internal and external stakeholders then use these data and information as inputs to make decisions, formulate evidence-based policies, and prepare PRSPs and pro-poor, anti-poverty programs. The data and information are for example used in the calculation of the proportion of people below national and international poverty datum

lines in the various reports such as the human development index (HDI) and MDGs prepared by external stakeholders like the WB and the different UN agencies.

The main objective of the research stated above is to assess and describe the information flow and activities performed specifically in the CSA of Ethiopia for producing data and information relating to poverty. This enables stakeholders to have a deeper understanding of players, data generation process, and the flow of data and information. It also aims to assess data quality of surveys in the CSA of Ethiopia using the IMF's DQAF and how they can be maintained and improved. Other objectives include:

- Assessing how the different stakeholders (researchers, government and NGOs, international institutions, the UN) use the information produced by the Central Statistical Agency of Ethiopia.
- Discussing the different views of poverty under different approaches such as the capability-based or the income-based approach, etc.
- Discussing Sen's[12] informational-based approach for capturing and assessing poverty and.
- Discussing the different informational bases (or indicators) under different initiatives and activities (such as MDG, human development, etc) and their relevance to poverty reduction.
- Discussing what effect does the quality of data and information used as informational basis (or the construction of indicators and poverty dimensions) has in combating poverty.
- Discuss whether mobile data collection improves the quality of data and information (the informational basis) or not and,
- Assessing which data quality dimensions of the DQAF of the IMF in the different statistical stages can be improved.

[12] Amarya Sen is winner of the 1998 Nobel Prize in Economics. He is known for his famous approach to development known as the capability approach.

1.6 Structure of the Study

This study is divided into seven chapters, with the Introduction as Chapter One. The second chapter discusses the choice of topic and the methodology used in the research; the third chapter deals with the concept of poverty from global, regional and national perspectives and discusses the role of NSOs in relation to the national poverty reduction strategy, in the fourth chapter the role of ICTs in NSOs in the various stages of the statistical process in relation to data quality is discussed. The fifth chapter outlines the roles of CSA in the PRSPs, its relation with external stakeholders, the ICT infrastructure and how it is used to collect, process, store, archive, and disseminate socio-economic data and information from household and establishment surveys, and, based on the evidence obtained from CSA through observation and practical experience, the sixth chapter discusses in detail, with flow charts and descriptions, how this data is generated. Inputs, processes, outputs, players, ICTs, and other resources required prior to data collection, data collection itself and post-data collection are also discussed. Conclusion and some preliminary recommendations are made in chapter seven.

1.7 Conclusion

In this chapter, the research problem, questions, and objectives are outlined. The research is premised on the idea that preserving and improving the quality of data and information produced by data-producing agencies such as the national statistical offices play a vital role in the global initiative launched to reduce poverty.. The questions stated above are addressed by studying the process or flow of data and information, the players involved and activities performed and the resources employed by the Central Statistical Agency of Ethiopia. The next chapter looks in more detail at the methodology adopted and argues for the particular suitability of an interpretive case study for carrying out this research.

CHAPTER TWO

RESEARCH METHODOLOGY

2.1 Introduction

This chapter is in five sections: the first discusses the methodology adopted for this research, reference materials used, and the reason for choosing the research topic; the second is a discussion of the relevance of the qualitative interpretive case study and its participant observation with documentation as a method of data collection. It discusses the importance of understanding the prevailing and complex relationships in generating data and information flow, the different players engaged and the resources employed to preserve and improve data quality at all stages of the statistical process. The reference materials used and choice of research topic are discussed in sections three and four respectively. The conclusions reached are presented in the last section.

2.2 Research Method

Research methods can be broadly classified into quantitative and qualitative. Qualitative research covers a plurality of research paradigms (positivist, interpretive and critical), within which there are many research methods (such as case studies, field studies, ethnography and action[13]. It takes place in the natural world, or uses multiple methods, and the researcher views social phenomena holistically[14] .Leedy also indicates that qualitative research focuses on phenomena that occur in the real world…and involve studying them in all their complexities[15]. The underlying aim of this study is to gain a deeper understanding of role players, the data generation process and the flow of data and information in the CSA of Ethiopia. As the objective of this study is to describe the relationship between the various ICTs used by NSOs and the role of NSOs in poverty reduction, it is important to understand

[13] Carroll & Swatman 2000, p.235
[14] Marshal & Rossman 1999, p.3
[15] Leedy & Ormrod 2005, p.133

how data and information are generated, the inter-relationships and interactions between the different stakeholders in the various statistical processes and the use of various ICTs.

A qualitative method with an interpretive philosophical perspective is used in this study. The interpretive approach in the information systems field is widely used: Orlikowski and Baroudi[16] state that ontologically (i.e. in terms of physical and social reality) interpretive information systems research assumes that the social world – social relations, organizations, division of labour – is not a *given*. Rather, it is created and reinforced by humans through their actions and interactions. In most countries nationally representative socio-economic data and information are produced by their respective national statistical offices. In the case of Ethiopia the production of data and information is the result of the interactions between the different internal and external stakeholders to CSA, ranging from government officials, international financial institutions to subject matter experts, and respondents, to mention only a few. The data and information are then used by these stakeholders in planning economic policies, developing PRSPs, monitoring and evaluation, etc. Understanding how the different players are engaged in their various activities and how they interact to generate data and information, enables the CSA not only to release of reports to stakeholders timeously, but also to ensure that the quality of data used to generate these information can enhance the decision-making process as well as the development of evidence-based anti-poverty programs and their monitoring. Premising on the epistemological basis of the interpretive philosophy (i.e. based on the nature of knowledge), Rosen suggests that understanding social processes involves getting inside the world of those generating it[17]. It is assumed that it is difficult to separate facts from values and that scientific knowledge is composed of both, unlike the positivist view that sees facts and values as distinct and that facts alone constitute scientific knowledge[18]. It is therefore important to look analytically at data-producing agencies such as the CSA of Ethiopia to understand and describe the different statistical processes and stakeholders involved in various surveys conducted by it.

[16] Orlikowski & Baroudi 1991, p.14
[17] as cited in Orlikowski & Baroudi 1991, p.14
[18] Walsham 1995, p.76

Colin[19] notes that case studies can be done on a group, on an institution, on a neighbourhood, on innovation, on a decision, on a service, on a program, and on many other things with a number of possible foci such as *best practices, policy implementation and evaluation, industrial relations, management and organizational issues, organizational cultures, processes of changes and adaptation; etc*. A case study in qualitative research provides an understanding of complex phenomena in real-time, with the opportunity to investigate and describe the process and interaction in organizations such as the CSA. Leedy[20] notes that in case studies extensive data is collected from the individual(s), program(s), or event(s) on which the investigation is focused. The author of this thesis, while employed by the CSA, used the experience acquired and observations made to collect data regarding the different statistical processes, the interactions between the different internal and external stakeholders, the different ICTs used to generate and process data and information and the CSA of Ethiopia. Ghauri and Grønhaug indicate that some of the data collection techniques in a case study may involve multiple sources such as verbal reports, personal interviews, and observations...financial reports, archives, and budget and operating statements, including market and competition reports as primary data sources[21]. Yin[22] also mentions the six sources of evidence for a case study as "*…documents, archival records, interviews, direct observation, participant observation,* and *physical artefacts".* It was noted by Walsham[23] that in the interpretive case studies method, interviews are the primary data sources for outside observers whereas interviews are still important data sources in the case of a participant observer or action researcher. Observation is therefore used as data-collection tool, with the author as the data collection instrument. Observation enables the researcher to collect as much data as possible about who is involved in what type of activity at any specific time and location. The research method used in this study includes case study and examination of various published materials within the interpretive philosophical approach.

Leedy states that methods of data collection in a case study include observation, interviews, and appropriate written documents and/or audio-visual material[24]. According to Collins,

[19] Collin 1993, p.146
[20] Leedy & Ormrod 2005, p.135
[21] Ghauri & Grønhaug 2005, pp.114-115
[22] as cited in Walsham 1995, p.78
[23] Walsham 1995. Interpretive case studies in IS research: nature and method, p.78.
[24]Leedy & Ormrod 2005, p.144

observations can be direct or participant. Direct observations are used if one wants to find out what people do, whereas in participant observation researchers assume a role or become part of the team in the day-to-day activities of an organization to observe the focus of enquiry such as the people, etc. He further states that with participant observation the primary data is derived from interpretations by the observer of what is going on and that the central research question is likely to be "*how?*"[25] Walsham again views participant observers or action researchers as being in a continuum in terms of their level of involvement, ranging from not being perceived by the people in the field as aligned with any particular individual or group within the organization, or only being concerned with making money as consultants or having as having strong prior views about specific people, systems or processes based on previous work in the organization, to being the fully and actively involved researcher, trying consciously and explicitly to change things in the way that they feel best[26]. Some of the advantages of participating in the observation method are that it allows the researcher to hear, see, and begin to experience reality to learn directly from his experience of the setting[27]. Ghauri and Grønhaug also state that collecting primary data through observation provides more objective and accurate data and information. The disadvantage is that its scope is limited, more costly and takes longer, while gaining access to the organization or the willingness of the respondents to deliver honest response may be problematic[28]. The author of this study believes that as an employee of the CSA of Ethiopia he has been more readily allowed to describe the units of analysis and collect data both at present and in the past.

In an interpretive case study, Orlikowski and Baroudi[29] argue that generalising from the setting (usually only one or a handful of field sites) to a population is not attempted; the intention is rather to understand the deeper structure of a phenomenon, which can then be used to inform other settings. Leedy also notes the weakness in a single case study, in that one cannot be sure that the findings are generalisable to other situations[30]. Leedy further states that when only a single case is studied, any generalizations made are naturally tentative and need to await further support from others studies – perhaps support from additional case

[25] Collins 1993, pp.190, 194-195
[26] Walsham 2006, p.321
[27] Marshal & Rossman 1999, p.106
[28] Ghauri & Grønhaug 2005, p.103
[29] Orlikowski & Baroudi 1991, p.5
[30] Leedy & Ormrod 2005, p.135

studies, other kinds of qualitative studies, or experimental research[31]. This study largely focuses on describing events, facts, and interactions taking place at the CSA of Ethiopia in an effort to generate data and information of use to internal and external stakeholders.

In order to answer the subsidiary and other related questions in this research as outlined earlier in Chapter One, participant observation is chosen as the method for data collection. Walsham pointed out that close involvement in research enables the researcher to have in-depth access to people, issues, and data[32]. Given the author's status as an employee of the CSA, it is possible to answer the *Who? Where? When? What?*, and *How?* questions, which are commonly asked in the case study research method.

In this research the author has in general attempted to observe the collection and flow of data, and the players involved. The locale of these observations was the Central Statistical Agency (CSA) of Ethiopia, both at the head office (between 1997 and December 2007) and at some field and branch offices (between 1997 and December 2003) during normal working hours. The activities and processes involved in the different surveys were observed at the head office in the period indicated above. These include questionnaire design, the training of trainers of survey data collection, data editing and coding, data entry, data cleaning, and tabulation. The participation (or rather the full involvement as employee of CSA) of the author ranged from providing inputs pertinent to data entry during questionnaire design by the subject matter specialist, to designing data entry forms for keyboard entry, preparing and implementing computer data cleaning and tabulation programs, and assisting data entry operators and data cleaning staff during the execution of the process. During the above period, at branch offices and in the field and while training enumerators and supervisors the following activities were observed: the different data collection steps, ranging from delineating the enumeration area (EA) to listing all households within the EA; the random selection of households from the listed households for interview; the interviewing itself; obtaining data like the dimensions of fields or plots, the measurement of food items consumed by households within a specific period of time; checking filled-in questionnaires for completeness and consistency, re-interviewing households and re-measuring areas by

[31] Leedy & Ormrod 2005, p.136
[32] Walsham 2006, P.321

supervisors to check for errors or falsifications by the enumerators. Different post-data collection activities observed included the developments of systems for data capturing, data cleaning, and tabulation.

Collin also notes that case studies are multi-method (typically involving observation, interviewing, and analysis of documents and records)[33]. Survey documents from CSA are available from their website and information regarding mode and method of data collection, data collection time for the specific surveys, etc was copied from the website and analyzed. These documents are found in the form of a meta-data, or survey report. As Collin indicates, a case study is not a survey, where reliability relies crucially on the characteristics of the data collection instrument; it relies on the trustworthiness of the human instrument (the researcher) rather than on the data collection techniques *per se*.[34] Leedy[35] also notes that the researchers should not confuse their actual observations with their interpretations of them. In this study, the analysis of the flows of data and information and how they are generated is informed by the author's actual work experience, both at the head office and the different branch offices, and is supported by the different internal documents from CSA and external documents from international organizations. Creswell[36] states that the analysis of cases studies from documents, occurrences, and other bits of data involves organization of specific "facts" in the case into logical order, categorizing them into meaningful groups in order to interpret them by examining the specific meanings they might have in relation to the case, then scrutinizing the interpretations and data for underlying patterns that can characterize the case more broadly, and finally constructing an overall portrait of the case to draw conclusions about possible implications. The author believes that the processes and players involved in the different activities in most of the surveys conducted by CSA are well described in this study and reflect what took place in the period specified above.

[33] Collin 1993, P.167
[34] Collin 1993, P.160
[35] Leedy & Ormrod 2005, p.146
[36] as cited in Leedy 2005, p.136

Methodology	An interpretive case study of the Central Statistical Agency of Ethiopia (CSA) following the qualitative method.
Data collection methods	
Observation	The author personally collected data at the head office and in the different branch offices in the field. Observations provided an understanding of the various statistical processes involving players and stakeholders on the one hand and the different ICTs on the other.
Documents	Internal documents such as published survey documents and the CSA's strategic plan were used. External documents included Government publications such as PRSPs, strategic plans as well as annual publications from international bodies like WB, IMF and the UN. Various related publications were also consulted
Method of analysis	
Observer impression and documents	The data collected through observation and from documents was structured into a coherent format and then analysed and interpreted

Table 0-1 Summary of research methodology

(Source: Author)

It is important to note that various international organizations rely on data and information from national statistical offices and other sources analysing poverty. They release reports and databases, devise anti-poverty programs, and conduct development activities in their respective countries. The quality of this data, its accuracy, reliability, timeliness and accessibility, etc. determines the value of the information and knowledge generated by other stakeholders using it. It is therefore important to study the process with all its activities, the players involved and the ICTs and human capital used at various stages to produce this data and information in order to preserve and improve its quality.

2.3 Major References

Both primary and secondary data sources from the CSA and other external sources are used for this research. The primary data was collected by means of personal observations at the CSA; the secondary sources included official national and international research publications,

reports and electronic materials related to this research. Below is a brief description of the internal and external documents consulted for this thesis.

The reference materials consulted include books and journals related to poverty reduction, KM and ICT. In addition, reports from organizations under the auspices of the UN such as the World Bank, United Nations Development Programme (UNDP), the UN and the International Telecommunications Union (ITU) were also used.

Two consecutive Ethiopian strategic plans prepared by the Ministry of Finance and Economic Development (MoFED) were also used. The ICT for Development (ICT4D) of Ethiopia as produced by the Ministry of Capacity Building in collaboration with the UNDP was also used as a reference.

The World Bank releases a number of periodical reports related to poverty. In addition to these reports such as World Development Indicators (WDI), a source book on poverty reduction strategy prepared by the World Bank was also consulted. The choice of this book as a reference is made because it is used by some countries as a guide in the development of poverty reduction strategies.

The UN, with the help of the different offices and specialized agencies under its auspices, also produced reports on the MDGs, which have been adopted as development frameworks by over 190 countries. These reports list eight goals, 20 targets, and over 60 indicators. Its 2008 report charts the progress made at global, regional and country level in respect of each goal. As part of the UN, the UNDP also releases its Human Development Report on a number of indices. Its 2008 report assesses about 177 UN member states for human development, using criteria such as life expectancy at birth, adult literacy rate, and GDP per capita.

The ITU, another specialized UN agency, releases various statistical reports measuring the progress and achievement of countries in different ICTs such as fixed and mobile cellular telephones, Internet penetration, and numbers of PCs, TVs, and radios. The ICT

Development Index (IDI) for example, is a report that charts progress in the development of ICTs in countries and monitors the so-called digital divide, (the ITU 2009, an Information Society: ICT Development Index). It covers over 150 countries and shows the progress made between 2000 and 2007, providing useful information to people who formulate policies and strategies. Access to and use of new technologies as part of the MDGs can be monitored by collecting data on ICT indicators, using either household or service provider surveys. Core lists of ICTs indicators are found in the manual prepared by ITU (the Manual for Measuring ICT Access and Use by Households and Individuals, 2009 Edition). The indicators in the manual are referenced in relation to other household surveys conducted by CSA such as the WMS and HICES, and the MDGs.

The reference material for Ethiopia is the Poverty Reduction Strategy Papers (PRSPs) prepared by MoFED. The first phase of the PRSPs is the Sustainable Development and Poverty Reduction Paper (SDPRP), covering the three years from 2003/04 to 2004/05, and Ethiopia's strategic framework for the five-year period 2005/06-2009/10 called the Plan for Accelerated and Sustained Development to End Poverty (PASDEP). In addition to this, the CSA's strategic plan and different survey documents have been used.

2.4 Rationale for the Research Topic

The percentage of people living under poverty line[37] in the sub-Saharan regions of Africa has increased since the 1980s. Most countries in this region are implementing poverty reduction strategies with the support from some UN agencies and financial institutions like the World Bank and International Monetary Fund. The Ethiopian government implemented the first phase of the Poverty Reduction Strategy Paper (PRSP) and it is now implementing the second phase of the PRSP, the PASDEP. According to the estimates of the Ethiopian Ministry of Finance and Economic Development, the proportion of poor or people living under the national poverty line in 2004/2005 was 38.7% at national level[38]. The Central Statistical Agency of Ethiopia (CSA) under the auspices of MoFED provides the inputs used in analysis aimed at policy formulation, monitoring and evaluation, and appropriate interventions. The

[37]Minimum expenditure (or its cost estimate) required by an individual to fulfil his or her basic food and non-food needs (Introduction to poverty analysis, World Bank 2005).
[38] MoFED, 2006, p.23

quality of the data is essential and preserving and improving data quality should be a core activity of the national statistical offices of any country. Finding out where in the statistical process specific types of ICTs can maintain and improve the quality of data and information is also important. Different macro-data and micro-data quality measurement standards as used by different stakeholders are proposed. Some of these quality dimensions relate to elements such as accuracy, reliability, relevance, timeliness, etc. The generation of quality data with all its dimensional elements depends on the type of ICT employed, its usage and the processes and systems applied at different levels.

It is hoped that this study will be of value to all national statistical offices and other data producing agencies as well as to the various stakeholders working in NGOs, other government offices and international organizations involved in poverty-related and development activities. One of its aims is to enrich the understanding of stakeholders (data users in government offices, NGOs, international and regional organizations) by describing the underlying process, the flow and production of data and information and the interactions between the different players engaged in statistical processes in national statistical offices in general and in the Central Statistical Agency (CSA) of Ethiopia in particular. This research aims to provide the insights that could help data and information users to understand information flow in the CSA of Ethiopia and the complex underlying processes.

2.5 Conclusion

In this chapter, the choice of the interpretive case study method is studied to show how data and information are produced. As mentioned above, the researcher (a participant observer) was the main collection agent. Observations, books, articles, and documents (internal and external) were also used as sources of data. It is assumed that with this method it is possible to describe the interactions between different internal and external stakeholders along with the generation and flow of data and information, from one stakeholder to another, internally (data collectors, programmers, subject matter specialists) and externally (between respondents, researchers, regional governments, international and regional organizations), during the pre-data collection, the data collection, and the post-data collection stages. Poverty-related data obtained from respondents in rural and urban areas passes through a

number of steps; it needs to be edited and coded, captured and processed before it can be effectively used in poverty analysis by external stakeholders for formulating evidence-based policies and strategies to be implemented in rural and urban households in order to reduce poverty. It is hoped that understanding the different views of poverty by different countries, international organizations, and scholars is essential in order to be able to use appropriate ICTs to improve data quality. Books and documents are the sources of data on poverty, ICTs and data quality. They are then combined with the observation made at CSA to have a better understanding of how, where, when, and what ICTs can be used to generate poverty-related data and improve its quality. Therefore the inclusion of consecutive chapters on poverty and ICT is important. The next chapter gives an overview of the concept of poverty and its trends. It also describes the chain of players and activities involved in preserving and improving the quality of data and information obtained from data-producing agencies.

CHAPTER THREE

POVERTY AND POVERTY REDUCTION

3.1 Introduction

This chapter examines the concept of poverty and the role of national statistical offices in global and national poverty-reduction initiatives and how poverty is perceived and measured by different countries, organizations, and scholars. Trends in poverty over a number of years showing the percentage of people under poverty are discussed, as well as how external stakeholders use the data and information gathered from NSOs and other line-ministries in their reports. The concept and type of indicators that are used by different stakeholders are described, along with the different types of data, indicators and information that are obtained from various data producing agencies. The way these are used in the operations of the World Bank in poverty assessment, discussion and formulation of policies and strategies, and implementation of lending programs is described here, as are Poverty Reduction Strategy Papers (PRSPs). The chapter concludes by stressing the importance of preserving and improving the quality of data and information obtained from data producing agencies.

3.2 Overview of Poverty

Poverty has a different face in every country and strategies used to reduce or eradicate poverty differ accordingly.. Regardless of the political or ideological orientation of developed or poor countries, there are always people who do not earn enough to provide them with food or all the other needs. Poverty may also mean being deprived of the social or political rights that would allow the poor to participate in income-generating daily activities.

Klugman defines poverty as a condition where households or individuals do not possess the resources or abilities to meet their current needs[39]. This definition compares the incomes and

[39] Klugman 2002, p.29

other attributes of individuals and uses a cut-off point (called a poverty line) below which individuals are categorized as poor or regarded as living in poverty. Most of the time poverty is categorized as either absolute or relative[40]. According to Dixon and Macarov[41] most high-income countries focus on the definition of relative poverty, whereas the absolute definition of poverty is applied to most poor or developing countries. According to Klugman[42]

"*Relative poverty lines are defined in relation to the overall distribution of income or consumption in a country; for example, the poverty line could be set at 50 percent of the country's mean income or consumption*, *whereas absolute poverty lines are anchored in some absolute standard of what households should be able to count on in order to meet their basic needs, using the food-energy intake method or the cost of basic needs method*. "

Ravallion[43] also says that "*absolute poverty has fixed 'real value' over time and space, while a relative poverty line rises with average expenditure"*.

The measurement of poverty based on income or a bundle of commodities draws criticism from various scholars. One of the most prominent of these, Amartya Sen[44], argues that categorising people solely on income or a bundle of goods across time and different geographical locations may exclude other important human deprivations such as the capability of individuals to enjoy better education, health, economic and political entitlements, etc. Sen[45] also argues that it is difficult to make poverty comparisons between individuals (even within a family) who may have the same demands for income and a commodity bundle such as same amount of a certain type of grain, or vegetable, or fruit, or similar household assets (such as a vehicle, housing, computer or Internet, TV, radio, air conditioner, etc) but may not enjoy the same utility, well-being, or quality of life. Some of these factors, according to Sen[46], could be personal differences (for example differences in age, gender, disability that demand different incomes), or environmental diversity (for example people who live in hot, cold, or disease-prone areas may require different levels of income to others who do not), or the allocation of one's income to changing social conditions

[40] Klugman 2002; Dixon 1998, Ravallion 1998
[41] Dixon & Macarov 1998, pp.1-2
[42] Klugman 2002, p.33
[43] Ravallion 1998, p.5
[44] Sen 1999
[45] Sen 1999, pp.69-70
[46] Sen 1999, p70-71

(such as public schooling, crime, epidemiology, pollution, community relationship, etc), or the exclusion of individuals whose income is relatively low in a high-income country, or the different distribution rules that apply to household members for the income earned by a family to attain goals related to particular perceived needs (such as sickness, age, or gender). Sen defines poverty as *deprivation of basic capabilities*[47]. His approach towards the assessment of poverty is widely known as the capability approach. This approach uses the terms *functioning*, *capability*, and *agency*. Capabilities are the freedom to enjoy functioning (such as being healthy and well-nourished, being safe, being educated, having a good job, etc), and agency refers to the ability to pursue goals through someone who acts and bring about change[48]. Capabilities are also defined as the substantive freedoms a person "*…enjoys to lead the kind of life he or she has reason to value."* as discussed below[49]. This definition differs from income poverty in that income poverty focuses on the assessment of basic needs (like housing, clothing, food, water, etc), other commodities, or household and personal assets in terms of income. Sen's definition also differs from the utilitarian approach to poverty, which focuses on the measurement and maximization of a person's happiness or the fulfilment of a person's desires. The particular assessment of poverty according to the various approaches depends on what informational base[50] each view takes as fundamental. The capability approach enhances the understanding of the nature and causes of poverty and deprivation by shifting primary attention away from means (in particular income) to ends that people have reason to pursue, and accordingly to the freedoms to be able to satisfy these ends[51].

In the US, the Census Bureau uses a set of income thresholds that vary according to family size and composition to determine who is in a state of poverty. If a family's total income is less than the family's threshold, then that family and every individual in it is considered to be impoverished. The official poverty thresholds do not vary geographically, but are adjusted for inflation according to the Consumer Price Index (CPI-U). The official poverty definition

[47] Sen 1999, p.87
[48] Deneulin & Shahani 2009, pp.9-10; Sen 1999, p.19
[49] Sen 1999, p.87
[50] The information that is needed for making judgements (Sen 1999, p.56)
[51] Sen 1999, p.90

uses monetary income before taxes and does not include capital gains or non-cash benefits (such as public housing, Medicaid and food stamps)[52]

In the UK the government categorizes people as poor if their income falls below a certain threshold of median income. The Department for Work and Pensions (DWP) is responsible for producing national statistics on Households Below Average Income (HBAI). According to the DWP (DWP 2009)[53] the thresholds of disposable income indicators (called equivalised income adjusted for household size) take housing costs into account, i.e. Before Housing Costs (BHC)[54] and After Housing Costs (AHC)[55]. The thresholds are either relative or fixed: a relative threshold is related to the contemporary median for each year's survey; a fixed threshold uses the median from a so-called anchor year, which is adjusted for inflation as appropriate. For example, the absolute threshold of 60 *percent of the 1998/99 median income* in 1998/99 is the same as the relative threshold, but the corresponding value in 2007/08 has been adjusted for inflation over the period (DWP 2009, p.175)[56]

In 1984 the European Union[57] classified the poor as persons, families and groups of persons whose resources (material, cultural and social) are so limited as to exclude them from the minimum acceptable way of life in the Member State in which they live". Social exclusions may result from factors as wide-ranging as unemployment, access to education, childcare and healthcare facilities, living conditions, as well as social participation[58]. According to the definition of poverty by Eurostat, individuals and families are categorized as poor not only when the equivalised income falls below the threshold of 60% of the national equivalised median income; when they are excluded from participating in social activities that may allow

[52] http://www.census.gov/hhes/www/poverty/definitions.html
[53] http://research.dwp.gov.uk/asd/hbai/hbai2008/contents.asp
[54] BHC includes social security benefits, net loss or profit from self-employment, net employment earning, income from private or occupational pensions, income from educational grants, etc. (see http://research.dwp.gov.uk/asd/hbai/hbai2005/pdf_files/appendices/appendix_1_hbai06.pdf)
[55] AHC is derived by deducting a measure of housing costs from total income measure (see above link)
[56] http://research.dwp.gov.uk/asd/hbai/hbai2008/contents.asp
[57] as cited in Spiker, Leguizamon & Gordon 2007, p.71
[58] European Commission 2010, p.63

them to lead a healthy life they are also regarded as being in poverty. However one concern about the use of the relative poverty line, according to Alkire and Santos[59], is that

"*…if all incomes in a society suddenly dropped sharply by exactly the same extent, leaving most of the society in a destitute situation from an absolute point of view, relative poverty would not change.*".

The World Bank describes poverty as follows:

> *"Poverty is hunger. Poverty is lack of shelter. Poverty is being sick and not being able to see a doctor. Poverty is not being able to go to school and not knowing how to read. Poverty is not having a job, fear for the future, living one day at a time. Poverty is losing a child to illness brought about by unclean water. Poverty often has a woman's face. Poverty is powerlessness, lack of representation and freedom."*[60]

The above view of poverty by the World Bank seems to have been influenced by the various works of Amartya Sen and others. Sen[61] views poverty broadly as a multi-dimensional phenomenon in terms of "*…deprivation of capabilities rather than merely lowness of incomes."* He argues that freedom can be seen "*…in the form of individual capabilities to do things that a person has reason to value."*[62] Alkire and Deneulin[63] explain the phrase *reason to value* to be an acknowledgment "*…that given our disagreements we do need to make some social choices"* regarding certain issues, for example those about discrimination, female genital circumcision, domestic violence, etc. According to Sen the five types of instrumental freedoms (political freedoms, economic facilities, social opportunities, transparency guarantees, and protective security) improve the capabilities of individuals[64]. Some of the advantages of these substantive freedoms support one another and include the right to participate in politics and criticize elected officials, the right to have freedom of uncensored expression, the right to participate in business in order to benefit from the country's economic growth, or the ability to participate in the various social institutions such as education and

[59] Deneulin & Shahani 2009, p.134
[60] World Bank.Retrieved on Jan 10, 2010 from http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTPOVERTY/0,,contentMDK:20153855~menuPK:373757~pagePK:148956~piPK:216618~theSitePK:336992,00.html#new_directions,
[61] Sen 1999, p.87
[62] Sen 1999, p.56
[63] Deneulin & Shahani 2009, p.10
[64] Sen 1999, pp.8, 38

health facilities so that one's productivity is improved as a result of a healthy life, or the right to be protected and get proper assistance regardless of background. Therefore, the well-being of individuals can be enhanced by expanding these substantive and primary freedoms such as freedom from starvation and under-nourishment, escapable morbidity and premature mortality, freedom to be literate and numerate, freedom of political participation and uncensored speech, etc. In other words development is seen as "…*the process of removing unfreedoms and of extending the substantive freedoms of different types that people have reason to value.".*[65]

While acknowledging that higher income is very important in achieving some of these capabilities, Sen contends that improved capabilities lead individuals to earn higher incomes.[66] Focusing solely on income inequality or income poverty as seen in the analysis of economic development of some countries and organizations, may lead to the avoidance and exclusion of other inequalities such as *"unemployment, ill health, lack of education, and social exclusion in the discussions and formulation of policies".*[67] The indicators used in evaluating the development of a country and the well-being of individuals should therefore not be limited to monetary terms such as GNP, GDP, etc, but should include the social, economic, and political factors on which capabilities depend. He also argues that although deprivations in income and capability correlate, this does not mean that income tells us a lot about capability – a broader *informational base* that integrates income and wealth with success and deprivation (Explain? HF) is needed in order to understand the poverty of human lives[68]. Ravallion[69] also indicates in his working paper that the capabilities and utility approaches are not fundamentally different and that the idea of *capabilities* as a welfare indicator is no substitute for utility (or some monetary measure of utility) as individual welfare indicator, but rather complements it. Increased income may therefore complement other capabilities of individuals and enable them to participate in other socio-economic activities in rural communities such as the right to own land and buy more agricultural inputs, or to have access to various social opportunities such as health and education, or to participate in the life of the community.

[65] Sen 1999, p.86
[66] Sen 1999
[67] Sen 1999, pp.107-108
[68] Sen 1999, p20
[69] Ravallion 1998, p.9

The author of this thesis shares the view that poverty is multi-dimensional in nature, as seen by Sen. This multi-dimensional nature as expressed in the above description is reflected in the social, economic, and political lives of people. Poverty is not only viewed as purely an incidental but also an ideological constraint.

Countries may also underestimate or overestimate the extent and causes of poverty. This applies as much to the working domestic policies that governments devise and preach during elections, as to international politics that may influence governments to change their policies. A broader informational base, approach and participation are required in order to assess human development, inform policies, and develop anti-poverty programs. Participatory poverty assessment (PPA) is one such tool that provides a better understanding of poverty.

Norton et al define PPA as

"*An instrument for including poor people's views in the analysis of poverty and the formulation of strategies to reduce it through public policy"*[70] This enables the poor to express their views about poverty. It also enables specific types of information about poverty and vulnerability to be noted and engages citizens in *debates about the best way of implementing poverty reduction policies and to assist the prioritisation of resources directed to pro-poor efforts*[71] among others. This type of qualitative information improves the design of the PRSPs and plays a vital role in the inclusion and exclusion of data variables or the construction of the data set that is going to be used in the complementary quantitative surveys such as the HICES and WMS.

Different countries in the developing world have home-grown policies, which they believe can bring about human and economic development. The World Bank encourages countries to develop their own policies and poverty-reduction strategies in collaboration with different stakeholders and which includes the technical assistance provided by its staff and NGOs as stated above. On the preparation and ownership of the PRSP, Klugman states that "…*the*

[70]as cited in Ellis & Woldehanna 2005, p.1
[71] Ibid.

strategy should be prepared by the government through a country-driven process including broad participations that promotes country ownership and its implementation."[72] However, there is criticism that some of the stakeholders such as Donors, the WB and IMF have been acting as promoters of the vested interests of some countries or multinational corporations linked to them[73]. Development assistance provided to poor countries is thus accompanied by pre-conditions[74] that require governments to institute changes to their economic policies. In the case of Ethiopia for example there is pressure from such institutions and some donors for the Government to liberalize and privatize the telecommunications, airlines and electric power sectors. The Government on the other hand believes that the expansion of infrastructures[75], for example telecommunications to rural households, can be done better if the sector remains under public rather than private ownership. The expansion of the telecom infrastructure enhances the capability of other government agencies involved in poverty reduction: the CSA as a government agency, for example, may improve its service delivery by collecting, processing, and disseminating poverty-related data and information through the use of better and more appropriate ICTs.

Understanding more about the extent and causes of poverty enables the State to prepare and implement appropriate anti-poverty programs. The State, as one of the primary stakeholders, uses the poverty-related data from NSOs as input in poverty analysis and in PRSPs. The relationship between the NSO's other stakeholders such as donors, international institutions and some UN specialized agencies can be seen as mutual as they all use the data collected by NSOs for their own consumption. Because of this mutual benefit, these stakeholders assist countries in developing financially and technically despite the complexity of some of the surveys and the capacity constraints of the NSOs. Some surveys are conducted with the support of donors and the WB. In the case of Ethiopia the periodic HICES and WMS were conducted in collaboration with the welfare monitoring unit of MoFED. Further support was provided by the WB, the Norwegian Trust Fund and development assistant groups (DAG)[76]. Similar surveys are conducted by NSOs of different countries. The contents of the questionnaires are similar, with some modifications to suit countries' varying needs.

[72] Klugman 2002, p.2
[73] http://en.ethiopianreporter.com/index.php?option=com_content&task=view&id=2274&Itemid=1
[74] Cheru, F. 2007, p.371
[75] http://www.bloomberg.com/apps/news?pid=newsarchive&sid=adot_hJ.1pyc
[76] CSA 2004b, p.3

3.3 Trends of Poverty

The World Bank's annual reports are one of the leading global development indicators and contain figures on poverty and the number of poor. The percentage distribution of the poor is computed from income surveys of households for countries or regions, i.e. the percentage of poor below a certain threshold such as $1, $1.25 or $2 a day. According to the World Bank's poverty estimates released in its World Development Indicators 2009 report[77], about 1.4 billion people in the developing world were living on less than $1.25 a day in 2005, down from 1.9 billion in 1981. Although poverty shows a decline from a global perspective, the classification of less than $1.25 a day at regional level shows both encouraging and depressing trends. The most substantial decline in poverty occurred in East-Asia and the Pacific region (from 77.7 percent or 1072 million in 1981 to 17 percent or 316 million in 2005) with China recording the biggest decline (from 84 percent or 835 million in 1981 to 15.9 percent or 208 million in 2005) with modest reductions in South-East Asian poverty rates. In Sub-Saharan Africa however the number of poor people almost doubled from 388 million or 50.9 percent to 211 million or 53.4 percent in 2005.. The report divides the regions into lower income (less than $1.25) and middle income (less than $2). In the lower income regions in absolute terms South-East Asia still has the biggest number of people living under poverty (40.3 percent or 596 million) followed by Sub-Saharan Africa (50.9 percent or 388 million), and East-Asia and the Pacific region (16.8 percent or 316 million). In terms of progress East-Asia and Pacific region are reducing poverty faster and are well on the way to meeting the MDG goal of reducing poverty rates by half by 2015. They will soon overtake Latin America, the Caribbean, and South-East Asia if they keep accelerating their growth rate. In its WDI 2007 report, the WB indicates that countries in Sub-Saharan Africa face difficulty in meeting the MDG goal of halving the proportion of poor by 2015 (Goal One of the MDG) if the average poverty rate remains above 40 percent[78]. In its 2008 MDG report the UN indicated that although "*the goal of reducing absolute poverty by half is within reach for the world as a whole;…(but) unless urgent action is taken, the proportion of people in Sub-Saharan Africa living on less than $1 per day is unlikely to be reduced by the target of*

[77] World Bank 2009, P.69-70
[78] WB 2007, p.4

one half[79]. Situations like the global economic slowdown and food price rises in the latter half of the 2008 may reverse the reduction in poverty already achieved in Sub-Saharan Africa in particular, making it difficult to meet the goal of reducing extreme poverty by half by 2015.

These poverty estimates are also supported by the findings of Chen and Ravallion; "…*not only has Africa emerged in the 1990s as the region with the highest incidence of poverty; the level of poverty is also markedly higher than that found in other regions*[80]. Among the middle-income regions Central Asia and Eastern Europe show a rise in poverty headcounts in 2005 while Latin America and Middle East and North Africa showed a drop in poverty headcounts. The table below shows estimates of people who lived in poverty at regional levels between 1981 and 2005:[81]

[79] MDG Report 2008, p.4
[80] Chen and Ravallion 2004, p.24
[81] World Bank 2009, p.70

Regional poverty estimates									
Region	**1981**	**1984**	**1987**	**1990**	**1993**	**1996**	**1999**	**2002**	**2005**
People living on less than 2005 ppp $1.25 a day (millions)									
East Asia & Pacific	1,072	947	822	873	845	622	635	507	316
China	835	720	586	683	633	443	447	363	208
Europe & Central Asia	7	6	5	9	20	21	24	21	17
Latin America & Caribbean	47	59	56	49	46	53	55	56	45
Middle East & North Africa	14	12	12	10	10	10	11	10	11
South Asia	548	548	569	579	559	594	589	616	596
India	420	416	428	435	444	442	447	460	456
Sub-Saharan Africa	211	241	256	295	317	355	382	390	388
Total	1,898	1,812	1,721	1,816	1,797	1,657	1,696	1,600	1,373
Share of people living on less than 2005 ppp $1.25 a day (%)									
East Asia & Pacific	77.7	65.5	54.2	54.7	50.8	36	35.5	27.6	16.8
China	84	69.4	54	60.2	53.7	36.4	35.6	28.4	15.9
Europe & Central Asia	1.8	1.4	1.1	2.1	4.4	4.8	5.3	4.8	3.8
Latin America & Caribbean	12.9	15.3	13.7	11.3	10.1	10.9	10.9	10.7	8.2
Middle East & North Africa	7.9	6.1	5.7	4.3	4.1	4.1	4.2	3.6	3.6
South Asia	59.4	55.6	54.2	51.7	46.9	47.1	44.1	43.8	40.3
India	59.8	55.5	53.6	51.3	49.4	46.6	44.8	43.9	41.6
Sub-Saharan Africa	53.4	55.8	54.5	57.6	56.9	58.8	58.4	55	50.9
Total	52.2	47	42.1	42	39.5	34.7	33.9	30.7	25.3
People living on less than 2005 ppp $2 a day (millions)									
East Asia & Pacific	1,278	1,280	1238	1,273	1,262	1,108	1,105	954	728
China	972	963	907	961	926	792	770	655	473
Europe & Central Asia	35	28	25	31	47	55	66	55	41
Latin America & Caribbean	89	109	102	95	95	106	110	113	94
Middle East & North Africa	46	43	47	44	48	52	51	50	51
South Asia	799	836	881	926	950	1,008	1,030	1,083	1,091
India	609	635	669	702	735	757	783	813	827
Sub-Saharan Africa	291	325	348	390	423	471	508	535	555
Total	2,538	2,622	2,642	2,760	2,825	2,800	2,870	2,791	2,560
Share of people living on less than 2005 ppp $2 a day (%)									
East Asia & Pacific	92.6	88.5	81.5	79.8	75.8	64.1	61.8	51.9	38.6
China	97.8	92.9	83.7	84.6	78.6	65	61.4	51.1	36.3
Europe & Central Asia	8.7	6.8	5.9	7.1	10.8	12.4	14.9	12.5	9.2
Latin America & Caribbean	24.6	28.1	24.9	21.9	20.7	22	21.8	21.5	17.1
Middle East & North Africa	26.7	23	22.7	19.7	19.8	20.2	18.9	17.6	16.9
South Asia	86.5	84.8	83.9	82.7	79.7	79.8	77.2	77	73.9
India	86.6	84.8	83.8	82.6	81.7	79.8	78.4	77.5	75.6
Sub-Saharan Africa	73.8	75.5	74	76	75.9	77.9	77.6	75.6	72.9
Total	69.9	68.1	64.7	63.8	62	58.6	57.4	53.6	47.3

Table 0-1 Estimates of people who live under poverty at regional levels

(Source: World development indicators 2009)

In order to get some basis for interpreting this data, we will first discuss data from China, the USA and the UK before turning to Ethiopia. It will then become clear that the focus of attention shifts in each of these countries.

Nearly 20 percent of the world's population lives in China. According to the Asian Development Bank[82], income poverty was defined in 1993 by China "*based on the cost of a basic minimum subsistence package of food plus a proportionate amount for essential non-food items"*. The poverty datum line is calculated as the sum of the food and non-food expenditures with adjustment for inflation for other years, using the national rural retail price index. Income in monetary terms is calculated from "*a fixed food bundle that was considered enough to maintain a minimum daily caloric intake (2,100 kilocalories) necessary for subsistence"* whereas the income from the non-food expenditures (clothing, housing, communication, fuel, medicine, education, entertainment, and others) is calculated as *"a fixed percentage of food expenditures, using an Engel's coefficient (ratio of food expenditures to total expenditures) of 0.6*[83]".. With 627 yuan (at 2002 prices) as the official rural poverty line, around 28.2 million or 3 percent of the rural population were categorized as poor in 2002; whereas in urban areas with 1875 yuan used (for diagnostic reasons) as the cut-off point by the National Bureau of Statistics, 10.5 million or 3.1 percent residents of the urban population fell below the poverty line in 2000. In a different report by ADB[84], using 668 yuan per capita per year (in 2004) as the official poverty line, there were 26 million poor people in 2004. However Park and Wang[85] suggest that China's figure of only three percent of the rural population being poor in 2000 is an under-estimate due to what they believe are potential sources of bias such as price indices, food bundles, non-food expenditures, income and expenditure measurements, etc.

The United States has a population of around 301,621,157 based on the projection made in July 1, 2007[86]. The US Census Bureau's report of 2007 indicates that, 37.3 million people were living in poverty, up from 36.5 million in 2006. Poverty rates in 2007 were statistically unchanged for non-Hispanic Whites (8.2 percent), Blacks (24.5 percent), and Asians (10.2 percent) from 2006. The poverty rate for Hispanics increased to 21.5 percent in 2007, up from 20.6 percent in 2006. The poverty rate for children under 18 years old increased (18.0

[82]World Bank 2009, p.70
[83] ADB 2004, p.16
[84] ADB 2005, p.2
[85] Park and Wang 2001, pp.390-393
[86] See http://www.census.gov/popest/national/asrh/NC-EST2007/NC-EST2007-01.xls

percent in 2007, up from 17.4 percent in 2006), while it remained statistically unchanged for people 18 to 64 years old (10.9 percent) and people 65 and over (9.7 percent)[87]

In the UK, the poverty profile below 60 percent of contemporary median net disposable household income in 2007/08 indicates that there were 2.9 million and 4 million children, 5.6 million and 7.5 million working-age adults, and 2.5 million and 2 million pensioners below the relative poverty indicator definition, i.e. before housing costs (BHC) and after housing costs (AHC) respectively[88]. The BBC reports that the signs of poverty in the UK include not having a High Street bank account, having to spend more than 10% of income on energy bills, and poor access to transport, employment opportunities or healthy food.[89]

In Ethiopia, like many other countries, poverty datum lines are measured and analyzed in order to assess policies and strategies and to design interventions. Here people below a certain poverty threshold are also considered to be poor – this line is measured by using the income and non-income dimensions method, based on the cost of basic needs[90]. The poverty lines were measured, using the 1995/96 national average constant prices. The datum line was taken as ETB 1075[91] per adult per year or 2200 Kcal[92] per day per adult, with the food poverty line constituting ETB 647.81 per adult per year.

According to MoFED the indices being used are: the poverty head count ratio, which measures the number of poor people below the poverty line; the poverty gap index, which provides information on how far a household's average per adult income/consumption is from the poverty line, and the poverty severity (squared poverty gap index), which takes into account not only the distance separating the poor from the poverty line (the poverty gap), but

[87] http://www.census.gov/hhes/www/poverty/poverty07/pov07hi.html
[88] DWP 2009, pp.90, 116, 114
[89] BBC, Tuesday, 26 July 2005, 13:20 GMT 14:20 UK, "The changing face of poverty". http://news.bbc.co.uk/2/hi/business/4070112.stm, retrieved on May 12
[90] This requires information on the prices of the goods that the poor consume. When price data is not available, researchers use the *food energy intake method*.
[91] ETB stands for the Ethiopian currency Ethiopian "*Birr*".
[92] 2100 Calories (2100 Kcal) per person per day is a standard widely used and proposed by the Food and Agricultural Organization (FAO) of the United Nations. "*The quantity of the bundle of food is determined in such a way that the bundle supplies the predetermined level of minimum caloric requirement (2,200 kilocalorie)... valued at local prices or at national prices*" (MoFED 2006, pp.21-22).

also inequality among the poor, so that a higher weighting is placed on households furthest from the poverty line[93]. Based on the Household Income and Expenditure Survey (HICES) conducted by CSA in 2004/05, the MoFED computed and compared the total national poverty, taking in to account both the food and non-food requirements (table 3-2 and table 3-3). It was found that the proportion of poor people below the national total poverty line has shown significant improvement in the five-year period, from 44.2% in 1999/00 to 38.7% in 2004/2005.

Level	**Indices**			**Changes in the Indices (%)**		
	1995/96	1999/00	2004/05	1999/00 over 1995/96	2004/05 Over 1995/96	2004/05 Over 1999/00
National						
Head count index (P0)	0.455	0.442	0.387	-2.7	-14.8	-12.4
Poverty gap index (P1)	0.129	0.119	0.083	-7.7	-35.4	-30
Poverty severity index (P2)	0.051	0.045	0.027	-12.2	-47.1	-39.8
Rural						
Head count index (P0)	0.475	0.454	0.393	-4.4	-17.1	-13.4
Poverty gap index (P1)	0.134	0.122	0.085	-8.9	-37	-30.8
Poverty severity index (P2)	0.053	0.046	0.027	-12.9	-48.3	-40.6
Urban						
Head count index (P0)	0.332	0.369	0.351	11.1	5.9	-4.7
Poverty gap index (P1)	0.099	0.101	0.077	2	-22.1	-23.6
Poverty severity index (P2)	0.041	0.039	0.026	-7.1	-38.2	-33.5

Table 0-2 Trends in total poverty indices at national level and by rural and urban areas.

(Source: MoFED 2006, Ethiopia: Building on Progress a Plan for Accelerated and Sustained Development to end Poverty (PASDEP, p.23))

Level	**Indices**			**Changes in the Indices (%)**		
	1995/96	1999/00	2004/2005	1999/00 over	2004/2005 over	2004/2005 over
National				1995/96	1995/96	1999/00
Head count index (P0)	0.495	0.419	0.38	-15.5	-23.3	-9.2
Poverty gap index (P1)	0.146	0.107	0.12	-26.8	-17.5	12.8
Poverty severity index (P2)	0.06	0.039	0.049	-34.494	-18.4	24.5

[93] MoFED 2006, p.22

Rural						
Head count index (P0)	0.516	0.411	0.385	-20.4	-25.5	-6.5
Poverty gap index (P1)	0.152	0.103	0.121	-31.9	-20.5	16.8
Poverty severity index (P2)	0.062	0.038	0.049	-39.2	-21.5	29
Urban						
Head count index (P0)	0.365	0.467	0.353	28	-3.3	-24.5
Poverty gap index (P1)	0.107	0.127	0.117	18.4	9	-8
Poverty severity index (P2)	0.044	0.047	0.048	6.8	8.4	1.5

Table 0-3 Trends in food poverty indices at national level and by rural and urban areas

(Source: MoFED 2006, Ethiopia: building on Progress: A Plan for Accelerated and Sustained Development to End Poverty (PASDEP, p.27))

The MoFED found that much of the decline in national poverty is due to the decline of poverty in rural areas, which in turn is the result of the implementation of a number of anti-poverty or pro-poor programs such as "*menu-based extension programs to support marketisation of small holder agriculture; the food security programs; and the recent productive safety net programs among others*"[94]. This result is also observed in the findings of Bresciani and Valdes[95], so that they conclude from their study that "*...in Chile the pro-poor role of agricultural expansion appears to be quite dramatic: agricultural growth tends to improve all measures of poverty significantly, with the head count falling between 8 and 11 percent as a consequence of a 4.5 percent increase in agricultural output; in India enhancing the ability of the poor to earn their living is the most sustainable way of poverty reduction and that agricultural growth over and above the population growth has a poverty reducing effect*. In the ten years (from 1995/96 to 2004/05) the depth and severity of total poverty in the urban areas of Ethiopia has decreased, as seen in the table above, going down from 9.9% to 7.7% and 4.1% to 2.6%; the proportion of poor increased from 33.2% to 35.1% however, whereas both the depth and severity in the food poverty indices have increased, with a minor drop in head count index.

The Participatory Poverty Assessment (PPA) is another survey conducted by the MoFED to supplement the WMS and HICES. Some of the aims of the PPA, which complement the quantitative surveys, are to provide insights into factors such as the causes and consequences of livelihood success or failure, , and to distinguish those policies under the SDPRP that are working well from those that have problems in their implementation, and to identify the types of improvement or service delivery valued by ordinary people, although these may sometimes differ from the priorities set by policymakers of the Central Government[96]. According to this report, females headed households in rural areas, while children, the landless and the elderly were the most prevalent vulnerable groups, with lack of land and livestock as the chief causative factors. Factors such as lack of education and skills, and an inability to start self-employment enterprises due to lack of savings or credit make youths, the disabled, the elderly, squatters and HIV/AIDS vulnerable, taking aspects of poverty such as

[94] MoFED 2006, p.25
[95] Bresciani & Valdes 2007, pp.78,129
[96]MOFED 2005. Can also be accessed from the University of East Anglia, *www.uea.ac.uk/dev/faculty/**Ellis**/research*

insufficient food for daily maintenance, crowded and unsanitary living conditions in poor quality housing, squatting in shacks made of plastic and wood, and exposure to personal danger into account.

Different stakeholders are involved in poverty reduction campaigns. They gather data and information from households, public and private enterprises and provide them to the government officials concerned so that they would be able to make evidence-based policies and track the progress made in the anti-poverty and other programs set out in the PRSP. The coordinated and concerted efforts of these stake holders (donors, World Bank, different government ministries, NGOs, etc) tackling the different dimensions of poverty require reliable information obtained from national statistical offices and other line ministries to devise and implement any effective intervention mechanism.

The WB, for example, works in close collaboration with different countries in the Sub-Saharan region in the fight against poverty. In its operational and policy manual, the WB indicates that the WB and IMF reviews the PRSPs of countries[97] in order to provide assistance to be outlined in the country assistance strategy paper (CAS). The CAS is prepared every four years, in consultation with governments through ministries and agencies, at a time when countries prepare their PRSPs[98]. In collaboration with concerned national institutes, the WB prepares poverty assessments of countries, based on the analytical outputs prepared by governments, the WB, donors, research organizations, etc and reflects all the findings obtained in the CAS. Based on the contents of the CAS (as one of the lending criteria) and the assessment of the country's policy and institutional framework, the WB decides on development policy lending[99]. Some of the data used in poverty assessment are obtained from national statistical offices (NSOs) through household surveys. The quality of this data and information (in terms of *availability* of clear data and information and meta-data, *periodicity*

[97] http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTPOVERTY/0,,contentMDK:20236175~menuPK:493904~pagePK:148956~piPK:216618~theSitePK:336992,00.html

[98] http://web.worldbank.org/WBSITE/EXTERNAL/PROJECTS/EXTPOLICIES/EXTOPMANUAL/0,,contentMDK:20064541~menuPK:64701771~pagePK:64709096~piPK:64709108~theSitePK:502184,00.html

[99] http://web.worldbank.org/WBSITE/EXTERNAL/PROJECTS/EXTPOLICIES/EXTOPMANUAL/0,,contentMDK:20240031~menuPK:4564187~pagePK:64709096~piPK:64709108~theSitePK:502184~isCURL:Y,00.html

and *timeliness*, *accuracy* and *reliability*, etc) therefore should be preserved. Regardless of the type of information base and approach to poverty being used, improving and preserving the quality of various socio-economic indicators constructed from household surveys obtained from NSOs and administrative records obtained from line ministries contributes much in assessing poverty. Data availability[100], which is one element in the *accessibility* dimension of the IMF's DAQF, is one of the most frequently cited data quality problems encountered in the assessment and comparison of poverty between different countries using various approaches mentioned in the previous section.

3.4 Poverty Indicators

Some of the indicators used by the UNDP to measure the human development index (HDI) of different countries include life expectancy at birth, adult literacy rate, combined gross enrolment ratio, and per capita GDP (PPP US$)[101]. Most of the MDGs indicators at global, regional and national levels are gathered from national statistical offices and different sector ministries. The UN's statistical division compiles them in a database for analysis to monitor progress made. It also reports the achievements of countries towards meeting the MDGs. The WMS and HIECS provide socio-economic indicators while indicators related to ICTs such as telephones and tariffs, the Internet, and computers are usually gathered from countries' telecommunication services providers. Some of the indicators such as total population, GDP, subscribers of fixed or mobile or Internet, etc are commonly used in the various reports.

3.4.1 What are Indicators?

Indicators are measures of quantitative and qualitative socio-economic and other aspects that enable governments and stakeholders to monitor and evaluate whether implemented policies have brought about the intended social, economic and other outcomes. Indicators can show the direction and gauge the level of improvement achieved. These achievements (such as in food security, life expectancy, gross domestic product, per capita income, gross enrolment ratio) can be shown in disaggregate forms such as by gender, age, urban-rural or geographical locations (at global, national, and different administrative levels within a country). Indicators

[100] See for example World Bank 2006, p.74; World Bank 2007, p.64; UN 2008, p.50; UN 2009, p.54.
[101] UNDP 2007, pp.225-227. Can also be accessed http://hdr.undp.org/en/humandev/learnmore/

can be computed as aggregates of variables or derived variables that convey some kind of information from panel or cross-sectional surveys at individual and household level. An indicator can also be defined as "*…a statistic which is taken to mean something else besides the core information it contains".*[102] Morse lists the following summary of characteristics of indicators used by managers and policy makers in selecting indicators[103]:

- *Specific – must clearly relate to outcomes that are being sought, although some also point out the advantages of indicators that reflect many more facets than those they directly measure (Kao and Liu, 1984)*
- *Measurable – implies that it must be a quantitative indicator, although "scores" or "ranks"' of more qualitative attributes are also acceptable.*
- *Usable – indicator must facilitate usage in terms of guiding management and/or policy.*
- *Sensitive – must readily change as circumstances change; ideally there should be a minimum time lag so that as circumstances change the indicator will also change.*
- *Available – it must be a relatively straightforward task to collect the necessary data for the indicator, even if it has to be collected (primary data) rather than being already available (secondary data). There are dangers here though. Care needs to be taken that the indicator does not become "data-driven" (Bayless and Bayless, 1982; Wish, 1986).*
- *Cost effective – it should not be expensive to obtain the necessary data.*

From a strategic point of view in reducing poverty, indicators are defined by Klugman[104] as "*…variables used to measure progress towards the goals where goals are the objectives a country or a society want to achieve*. Indicators are further classified by Klugman as intermediate (consisting of inputs and outputs) and final (consisting of outcomes and impacts) where the latter measures the progress towards the goals set and change slowly over time and

[102] cited in Midgely and Piachaud 1984: 39, Spicker et al P.106
[103] Morse 2004, p.30
[104] Klugman 2002, p.108

the former change more rapidly giving a chance to take corrective action during the implementation of a program. These terms are defined by Klugman as follows:[105]

> *When an indicator measures a factor that determines an outcome or contributes to the process of achieving an outcome, we call it an "input" or "output."*
>
> *For example: Input indicators could be a measure of public expenditure on class rooms and teachers' salaries. Output indicators could be measures of classrooms built and teachers trained.*
>
> *When an indicator measures an outcome or the effect of an intervention on individuals' well-being, we call it an "impact" or "outcome" indicator.*
>
> *Outcome indicators capture access to, use of, and satisfaction with public services. Impact Indicators measure key dimensions of well-being such as literacy and freedom from hunger.*
>
> *For example: Final goal or one dimension of well-being could be literacy. Outcome indicator could be a proportion of people of a certain age who can read a simple text and write their name*.

The above definition of indicators are widely used by a number of governments in developing countries that are implementing PRSPs. Types of indicators are best explained in the figure below from Klugman[106]:

[105] Klugman 2002, pp.108-110 and 164
[106] Klugman 2002, p.108

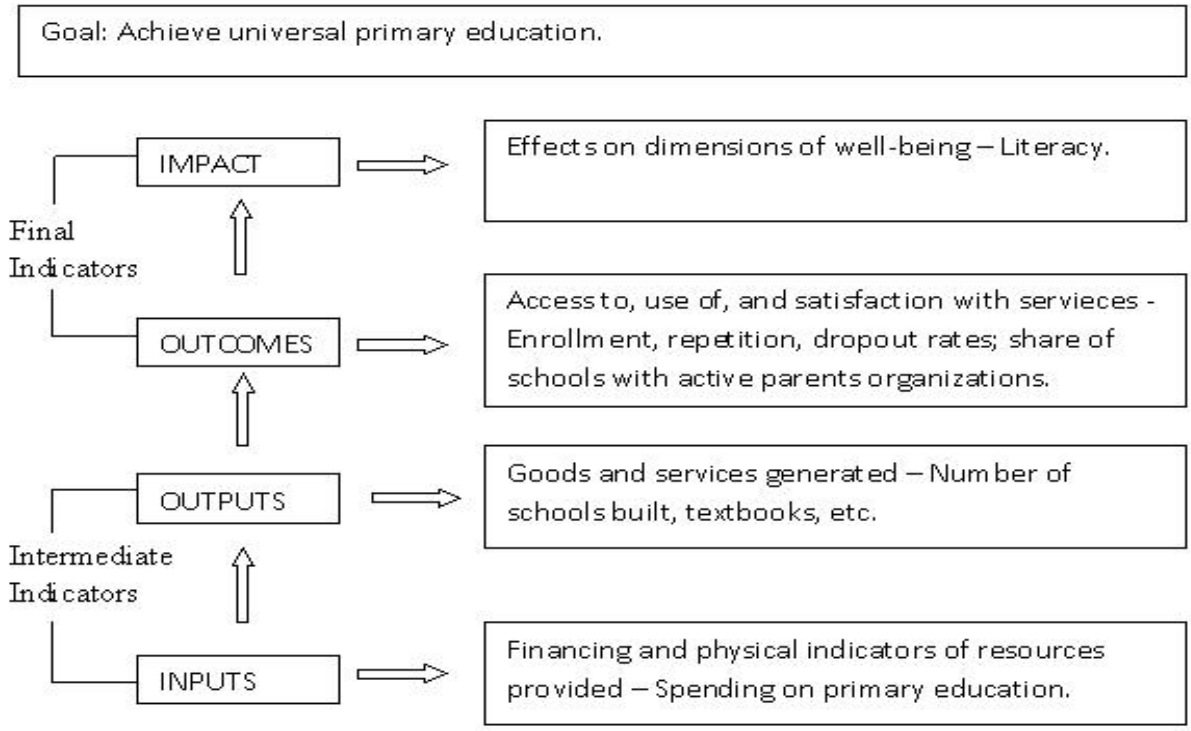


Figure 3-1 Intermediate and final indicators

(Source: Klugman, 2002, Source book of poverty reduction strategy, p.108)

Governments, International organizations, Donors, and NGOs use indicators in planning policies and prioritizing actions, either to implement or to allocate funds and assistance. Morse pointed out that indicators simplify complexity in order to act in organizations with processes either top-down or bottom-up "W*e need to know when we have achieved what we want to achieve"*[107].

3.4.2 MDGs Indicators

Originally the MDGs were developed as a framework consisting of eight goals, 18 targets and 48 indicators (Road Map towards the Implementation of the United Nations Millennium Declaration 2001). According to Hulme[108] the MDGs came into effect in an on-going process with no precise beginning and end, with one summit following another with the participation

[107] Morse 2004, p.31
[108] Hulme 2009

of a number of players (such as the Development Assistance Committee-DAC of the OECD, UN, IMF, WB, developed and developing countries, NGOs, and individuals). He states that the MDGs evolved by setting goals and targets in a number of world summits and conferences held in the 1990's. Among them are the World Conference on Education for All, the UNCTAD conference on the Least Developed Countries, a conference on drug problems, and the World Summit for Children (focusing on infant and maternal mortality, primary education, malnutrition, safe water and sanitary services), the UN Conference on Environment and Development (a.k.a. the "Earth Summit" or "Rio Summit"), the International Conference on Food and Nutrition in Rome, which set a target of halving the number of hungry people in the world, the World Conference on Human Rights in Vienna, the International Conference on Population and Development (ICPD) in Cairo, the World Summit on Social Development (WSSD) in Copenhagen, and the UN Fourth World Conference on Women in Beijing[109]. These were followed by a discussion to arrive at a set of common development goals in 2000 by the IMF, OECD, WB, and the UN, with international development goals (IDGs) of the OECD acting as basis in developing countries which eventually led to the approval of the Millennium Declaration with the support of 189 countries and 147 heads of states and governments[110]. This was followed by a more technical discussion on the selection and construction of indicators, the demotion of goals to indicator and target status (for example, demoting the safe drinking water goal to an indicator, promoting 100 million slum dwellers to a target under environmental sustainability goals), and the addition of a goal 8[111].

The official MDGs site of the UN states that "…*the Millennium Development Goals and targets come from the Millennium Declaration, signed by 189 countries, including 147 heads of State and Government, in September 2000 and from further agreement by member states at the 2005 World Summit*[112]. Later in 2007, the MDG monitoring framework was revised to include the four new targets agreed on by member states at the 2005 World Summit (see the official UN site for the MDGs indicators).

[109] Ibid., pp.8-9
[110] Ibid., pp.32-40
[111] Ibid., pp.40-41
[112] The official UN site for the MDGs indicators, http://mdgs.un.org/unsd/mdg/Host.aspx?Content=Indicators/OfficialList.htm, retrieved May 28, 2009.

The 2008 MDG report indicates that data used in the construction of MDG indicator is basically collected periodically from ministries and national statistical offices around the globe by concerned international agencies (such as FAO, UNESCO, WHO, ITU, WB, IMF, ILO, UNDP and UNICEF) and then compiled by the United Nations Inter-Agency and Expert Group on MDG Indicators[113]. The table in Appendix A shows the revised framework of goals, targets, and indicators of the MDG effective as of 15 January, 2008. Some of the indicators in the MDG used to measure progress include (for the full list see Appendix A):

- Proportion of population earning below $1 (PPP) per day
- Poverty gap ratio
- Share of poorest quintile in national consumption
- Growth rate of GDP per person employed
- Proportion of population below minimum level of dietary energy consumption
- Literacy rate of 15-24 year-olds, women and men
- Ratios of girls to boys in primary, secondary and tertiary education
- Under-five mortality rate
- Maternal mortality ratio
- HIV prevalence among population aged 15-24 years
- Proportion of population using an improved drinking water source

3.4.3 Human Development and Human Poverty Indicator

The UNDP has been preparing the Human Development Report (HDR) since 1990. It is comprised of various aspects of individuals in terms of their demographic characteristics, access to basic services and sanitation, mortality, in the areas of education such as public spending, enrolment, and technological diffusion. It also covers topics related to unemployment, gender, and income inequality. The capability approach proposed by Amartya Sen acted as a foundation for the successive HDRs, focusing on human capabilities rather than income alone. Alkire and Deneulin state that the intention of the HDR is "…*to assess the quality of life of a population and be an advocacy tool for its improvement …*

[113] UN 2008, p.50

draws on data regarding health, education, nutrition, work, political freedoms, security, the environment and many other aspects of people's lives … [with a] political purpose of raising awareness and generating debate on public issues and concerns which would otherwise not be on the political agenda."[114]. Some of Sen's works are now partially used by the UNDP's human development report (HDR) in some of its indicators such as the HDI and HPI. Anand and Sen state that "*…the HDI has played a role in broadening evaluative attention from such gross measures of economic progress as the gross national product (GNP) per head (or the average national income) to lines of analysis that are sensitive to the removal of deprivation of different kinds as well as expansion of opportunities in general."*[115]. According to the UNDP's world development report[116] the HDI "*…measures the average achievements in a country in three basic dimensions of human development*." The three dimensions (a long and healthy life, knowledge, and a decent standard of living) correspond to the five indicators (life expectancy at birth, adult literacy rate and gross enrolment ratio, and GDP per capita (PPP US$)).

The HPI, on the other hand, measures deprivations in the three basic dimensions of human development captured in the *HDI*[117]. The HPI has two versions HPI-1 and HPI-2. The HPI-1, which are used for developing countries has three dimensions (a long and healthy life, knowledge, and a decent standard of living) corresponding to four indicators (probability at birth of not surviving to age 40, adult illiteracy rate, percentage of population not using an improved water source and percentage of children under weight-for-age), and the HPI-2 which is used for selected OECD countries has four dimensions (a long and healthy life, knowledge, a decent standard of living, and social exclusion) corresponding to four indicators (probability at birth of not surviving to age 60, percentage of adults lacking functional literacy skills, percentage of people living below the poverty line, and long-term unemployment rate).

[114] Alkire & Deneulin 2009, p.24. Can also be accessed from International Development Research Centre, http://www.idcr.ca/en/ev-146685-201-1-DO_TOPIC.html
[115] Anand and Sen 1997, p.3
[116] UNDP 2009, p.356
[117] UNDP 2009, p.357

The measure of deprivations in HPI-1 and HPI-2 uses similar labels for the first three dimensions but the indicators used to construct the indices are different. In the long and healthy life dimension, for example, the HPI-1 sets the age in the life expectancy at birth indicator as 40 years, whereas in the HPI-2 it is set at 60 years. Anand and Sen in their technical note on the choice and construction of HPI[118] however argued that the choice of the indicators to be used in the human poverty index cannot be but sensitive to the context of the evaluation, and in particular to the characteristics of the countries for which this index is primarily intended. They also justify the appropriateness of using illiteracy rather than enrolment at different levels in the knowledge dimension by addressing the issue of *data quality* problems such as *availability*, *accessibility* and *reliability*[119]. Some deprivations such as illiteracy are more prevalent in developing countries than in the developed countries due to their social, economic and political situations. The assessment of poverty using HPI therefore takes these into consideration.

Both HDI and HPI are aggregate-based on the weightings attached to the indicators that constitute them. In many of the HDRs the indices are calculated using similar methods. For example in the 2007/2008 Human Development Report[120] the HPI-1 is computed as HPI-1 = $[1/3\ (p_1^{\alpha} + p_2^{\alpha} + p_3^{\alpha})]^{1/\alpha}$ where P1 is the probability at birth of not surviving to age 40, P2 is adult illiteracy rate, P3 is the unweighted average of percentage of population not using an improved water source and children under weight-for-age (i.e. ½(percentage of population not using an improved water source + percentage of children under weight-for-age), and α =3.

The 2008 HDR, for example, shows that Ethiopia ranked 169th with an HDI value of 0.389 out of 179 countries in the United Nation's Human Development Report (HDR) with life expectancy of 52.2 years, adult (15 years and over) literacy rate of 35.9 percent, combined GER in education of 45.1 percent, and GDP per capita (2006 PPP US$) of 700. It is classified in 2006 as a low HDI[121] country. In the Human Poverty Index (HPI-I) which excludes the high income OECD countries however, the report ranked Ethiopia 130th with the value of

[118] Anand and Sen 1997, p.6
[119] Anand and Sen 1997, p.8
[120] UNDP 2007/2008, p.357
[121] *Human development classifications:* high human development (with an HDI of 0.800 or above), medium human development (HDI of 0.500–0.799) and low human development (HDI of less than 0.500).

51.6 by using its measuring dimensions such as probability of not surviving to age 40 at 33.3%, adult literacy rate (proportion of population 15 years and above) at 64.1%, population not using an improved water source at 58%, the proportion of under-weight children aged under five years at 38%. It also uses percentage of population below different poverty lines as another dimension to measure HPI-I. These are based on recent data obtained between the years 2000 - 2006 for income poverty below 1.25 $ a day and 2 $ a day, and between the years 2000 - 2007 for the national poverty line as defined by national authorities. Accordingly the percentage of population in Ethiopia earning below 1.25 $ a day was 39%, below 2 $ a day was 77.5% at 2005 international prices adjusted for PPP, and below national poverty line 44.2%.

Below are some of the indicators available in the various publications of the HDR. The data used to construct the indicators is drawn from various data sources such as government agencies and UN specialized agencies:

- Probability at birth of not surviving to age 40 (or 60 % of cohort)
- Adult illiteracy rate (% aged 15 and older)
- Population not using improved water source (%)
- Children under-weight for age (% under age 5)
- Population below income poverty line (%, $1 a day $2 a day, national poverty line)
- Long term unemployment (as % of labour force)
- People lacking functional literacy skills (% of aged 16-45)
- Population below income poverty line (50% of median income, $11 a day, $4 a day)
- Total population (millions)
- Annual population growth rate (%)
- Total fertility rate (births per woman)
- Health expenditure (public % GDP, private % GDP, per capita PPP US$)
- One-year old immunization against (% tuberculosis, measles)

- Birth attended by skilled health personnel
- Physicians (per 100,000 people)
- Population using improved sanitation (%)
- Population using improved water source (%)
- Population undernourished (%)
- Children under-weight for age (% children under age 5)
- Children small for age (% children under age 5)
- Infants with low birth weight (%)
- Infant mortality rate (per 1,000 live births)
- Under five mortality rate (per 1,000 live births)
- Public expenditure on education (as % of GDP, as % of government expenditure)
- Adult literacy rate (% aged 15 and older)
- Net primary enrolment rate, net secondary enrolment rate (%)
- Tertiary students in science, engineering, manufacturing and construction (% of tertiary students)
- Telephone mainlines (per 1,000 people)
- Cellular subscribers (per 1,000 people)
- Internet users (per 1,000 people)
- Research and Development (R & D) expenditure (% GDP)
- Researchers in R & D (per million people)
- Female economic activity (aged 15 and older as % of male rate)
- Employment by economic activity - agriculture, industry, service ((% of men and women)
- Contributing family workers (% of men and women)
- Year women received right to vote, stand for election.
- Women in government at ministerial level (% of total)

- Seats in parliament held by women (% of total)

3.4.4 ICTs Indicators

The ITU is a specialized UN agency that collects telecommunications and ICTs related statistical data from telecommunications ministries, regulatory agencies, national statistical offices and industries related to ICTs. The data and information collected from administrative records cover broad topics related to services and ownership of telecommunications and ICTs while that from household surveys covers access to, proximity, and usage of ICTs such as telephones. This data and information are used by analysts, researchers, the Press, and countries. In some cases the ITU indicates either that data is not available or there are delays in getting both administrative and household data, mainly from developing countries. Below are some of the indicators listed in ITU's database (see Appendix B for the full list):

- Mobile cellular - price of 3-minute local call (off-peak - US$)
- Price of a 3-minute fixed telephone local call (off-peak rate, and peak rate)
- Price of a 3-minute fixed telephone local call (off-peak rate, peak rate - US$)
- Mobile cellular telephone subscribers (post-paid + pre-paid)
- Mobile cellular telephone subscribers (digital)
- Residential, and business monthly telephone subscriptions (US$)
- Mobile cellular monthly subscriptions
- Coverage of mobile cellular network (population, in %)
- Total revenue from all telecommunication services (US$)
- Main (fixed) telephone lines per 100 inhabitants
- Mobile cellular telephone subscribers, television receivers, and personal computers per 100 inhabitants
- Public payphones per 1,000 inhabitants
- Percentage of households with a radio
- Dial-up, DSL, and total fixed broadband Internet subscribers

3.4.5 WMS and HICES Indicators

According to the Ethiopian Central Statistical Agency, "…*the basic welfare indicator focuses on basic socio-economic indicators, which are vital inputs to the process of monitoring and evaluation of policies, particularly in poverty reduction strategies".*[122]. The non-income dimension of poverty is obtained from the WMS whereas the income-dimension is obtained from the HICES. The WMS provides data and indicators that give a general picture of basic needs of households and individuals reflecting their living standard. These basic needs include *education, health, child nutrition and care, access to and utilization of basic facilities, housing and housing amenities (drinking water, sanitation, energy, etc.), household assets, selected indicators of living standards, vulnerability (shocks and coping mechanisms, food security, etc.), and HIV/AIDS and basic population characteristics*[123].

Below is the list of basic welfare indicators (mostly aggregates) at different levels of disaggregation from the four rounds of surveys conducted in 1996, 1998, 2000, and 2004 by CSA and disseminated using publications in the form of maps, graphs and tables as well as using CD-ROMS and the Internet to stakeholders[124] (see CSA website):

1. Distribution of births within the past 5 years by place of delivery
2. Gross Enrolment ratio at primary school level, grade 1 to 6
3. Gross Enrolment ratio at primary level, grade 1 to 8
4. Gross enrolment ratio at secondary level
5. Households by type of cooking fuel
6. Households by sources of drinking water
7. Households distance in km to the nearest health service
8. Literacy rate age 10 years and above
9. Literacy rate in the last 4 years

[122]CSA, retrieved Feb 12, 2010, from http://www.csa.gov.et/index.php?option=com_content&view=article&id=81%3Abasic-_welfare-indicators&catid=63&Itemid=50
[123]CSA 2004b, p.5
[124]CSA, retrieved Feb 12, 2010, from http://www.csa.gov.et/index.php?option=com_content&view=article&id=81%3Abasic-welfare-indicators&catid=63&Itemid=50

10. Net enrolment ratio in primary school, grades 1 to 6
11. Net enrolment ratio in secondary school, grades 7 to 12
12. Population with health problems in the last two months
13. Prevalence of stunting, wasting and under-weight, ages 3 to 59 months
14. Prevalence of wasting, ages 3 to 59 months
15. Primary school drop-out rate
16. Total population size by sex
17. Total population by region.

3.5 Poverty Reduction Strategy Papers (PRSPs)

According to the IMF, the PRSP was conceived by the IMF and the World Bank in 1999, mainly for heavily indebted low-income countries so that governments in these countries could prepare their own poverty strategy papers in a participatory process together with different domestic and external stakeholders[125]. It is a national plan aimed at pulling the poor out of poverty and bringing sustainable economic growth. A PRSP describes the macroeconomic, structural and social policies and programs that a country will pursue over several years to promote broad-based growth and reduce poverty, as well as external financing needs and the associated sources of financing[126]

Klugman[127] suggests that these strategies should be:

- Country-driven and -owned, predicated on broad-based participatory processes for formulation, implementation, and outcome-based progress monitoring;
- Results-oriented, focusing on outcomes that would benefit the poor;

[125] IMF, retrieved 17April 17, 2009 from http://www.imf.org/external/NP/prsp/prsp.asp ,
[126] IMF, retrieved 17April 17, 2009 from http://www.imf.org/external/np/exr/facts/prsp.htm
[127] Klugman 2002, p.3

- Comprehensive in scope, recognizing the multidimensional nature of the causes of poverty and measures to attack it;
- Partnership-oriented, providing a basis for the active and coordinated participation of development partners (bilateral, multilateral, nongovernmental) in supporting country strategies; and
- Based on a medium- and long-term perspective for poverty reduction, recognizing that sustained poverty reduction cannot be achieved overnight.

Different countries have their own version of PRSP[128] (see website of the World Bank), Rwanda for example has PRSP (2002-2005) and Economic Development Poverty Reduction Strategy (EDPRS) (2008-2012). Nigeria has the National Economic Empowerment and Development Strategy (NEEDS) (2003–2007), and draft NEEDS 2 (2008-2011). Mozambique has PARPA I (2001-2005) and PARPA II (2005-2009), Tanzania has PRSP (2000/01 -02/03) and the National Strategy for Growth and Reduction of Poverty (NSGRP) (2005/06-2009/10). Zambia has PRSP (2002-2004), and Fifth National Development Plan (FNDP) (2006-2010). Most of these PRSPs are done in participatory process and the Ministry of Finance for each country normally takes the responsibility for coordinating the processes.

In Ethiopia the MoFED is in charge of this task. According to the MoFED[129] the PRSP in Ethiopia started in 2000 in collaboration with other stakeholders (donors) and evolved into the Sustainable Development and Poverty Reduction Program (SDPRP) in 2002, which covered the three-year period 2002/03-2004/05. The Plan for Accelerated and Sustained Development to End Poverty (PASDEP) represents the second phase of the PRSP which covers the five-year period 2005/06-2009/10. The PASDEP defines the national strategy for development, lays out directions, and outlines the major policies in the major sectors with the main objective of eradicating poverty within this time frame.

As stated above, different stakeholders take part in policy formulation, implementation, monitoring and evaluation, and intervention. In other words, a coordinated and well

[128]World Bank, http://web.worldbank.org/WBSITE/EXTERNAL/TOPICS/EXTPOVERTY/EXTPRS/0,,contentMDK:20195487~menuPK:384207~pagePK:148956~piPK:216618~theSitePK:384201,00.html
[129] MoFED 2006

organized mechanism between stakeholders such as different government ministries, NGOs, Community and faith based organizations, donors, UN, and financial institutions such as IMF and the World Bank, development agencies, and the people is required at the various stages of the process of preparing the PRSP and its implementation.

The role NSOs play as stakeholders is by collecting, processing, storing, and disseminating quality data and information to other stakeholders. For example, NGOs who are engaged in drilling for water may need information regarding, among others, household access and distance to the nearest clean water source so that they can prioritize which population group is the neediest in terms of potable water, and to provide the water sources in this community or village. The *availability* of such types of information from data producing agencies such as the NSOs and line ministries enhances the operations of other stakeholders who are engaged in poverty-oriented programs. Analysis of surveys such as the Welfare Monitoring Systems (WMS) which provides a non-income poverty dimension of poverty and access to basic services and their quality, the Household Income and Expenditure (HICES) which provides the income dimension of poverty, Labour Force, Consumer Price Index (CPI), and Food Security can provide indicators on poverty. Most of this data is collected by national statistical offices and in the case of Ethiopia by the CSA, while others come from various government line ministries and non-governmental organizations in the form of administrative records. In addition, qualitative participatory poverty assessment surveys are conducted by other institutions. Broadening the range of information on households and individuals regarding their socio-economic, demographic, agricultural, and political aspects is therefore essential in assessing human development and to formulate and make better evidence-based policies and decisions. In addition to this, the poverty assessment surveys can also assist in assessing policies and strategies that are already under implementation in the various ongoing projects.

3.6 Conclusion

In this chapter the multidimensional nature of poverty was discussed. It was shown that poverty is world-wide and that it is manifested in different faces and with different degrees of severity. It was also shown that designing and monitoring anti-poverty programs requires

different sets of indicators. Some countries use similar indicators such as poverty head count or poverty gap ratio to measure poverty but poverty datum lines may differ. It was also indicated that the view of poverty and the preparation of the PRSPs has some ideological orientation. Most high-income countries apply the relative definition of poverty, whereas low income countries apply the absolute definition of poverty. Sen argues that the capability approach to poverty has a broader informational base than other approaches such as the utilitarian or libertarian approaches. He states that "*…the broadening of the informational base from income to the basic capabilities enriches our understanding of inequality and poverty in quite radical ways.*"[130]. The ability of individuals to live without fear irrespective of their incomes, for example, is severely restricted in countries where there are different forms of unrest (ethnic, racial, religious, civil war, war with other countries, domestic violence, etc). Under the capability approach, poverty is assumed to be deprivation of basic capabilities such as the prevalence of high rates of mortality, illiteracy, malnourishment, unemployment, ill health, lack of education, social exclusion, etc[131]. One of the reasons is that under this approach the extent and cause of poverty can be assessed and it is possible to inform some policy issues for demanding actions.

In most countries the national statistical offices conduct nationally representative household and establishment surveys that are used to measure poverty and identify vulnerable groups. Accurate, reliable, and timely socio-economic data and Information are used by government officials and policy planners. They are important in monitoring and evaluation of anti-poverty programs and are used as evidence to make decisions. They enable policy makers to devise intervention programs and also enable the progress towards national and international initiatives such as the MDGs to be tracked. The effectiveness of these policies, strategies and various anti-poverty programs can therefore be affected by the quality of data and information among other things. This means that the quality of data and information produced by data-producing agencies needs to be preserved and improved. Improving data quality requires a coordinated and concerted effort by all stakeholders involved in the production of the data along with better and appropriate ICTs, and systems.

[130] Sen 1999, p.97
[131] Sen 1999, pp. 103,109

The role of statistical offices, as data producing agencies, is therefore undeniably important. An understanding of what data and information are required, how these are generated, who generates them, what resources are needed and how they are used, etc within the national statistical system in general and the national statistical office in particular at different levels of the statistical process, is also needed to meet the growing data quality demand of stakeholders. Coordination and collaboration with stakeholders, and application of appropriate and better ICTs are believed to be important factors that may preserve and improve some of the data quality dimensions stated in the DQAF. The next chapter discusses the different statistical processes, the types and usages of ICTS, and the different data quality dimensions that may be affected or improved at each stage.

CHAPTER FOUR

DATA QUALITY AND ICT

4.1 Introduction

This chapter starts with a discussion of the approach to and dimensions of data quality, mainly in relation to that collected by NSOs. The concept of data, information and knowledge is discussed next, followed by a discussion on the different activities performed in NSOs to generate data and the use of ICTs in the different stages of the statistical process namely before data collection, during data-collection, and post-data collection stages the fourth section. The discussion on the various modes of data collection draws on available literatures on mobile data collection using PDAs and laptops in relation to improving specific data quality dimensions. The chapter ends with the conclusions drawn.

4.2 Data Quality

Different national and regional statistical offices, international organizations, and other stakeholders have their own definitions and measurement of macro-data and micro-data quality. As part of standard quality procedures during the post-data collection stage, most NSOs use different computations such as a statistical coefficient of variation and other techniques to decide from its accuracy whether the collected data or information is fit for use. The quality of data noted by different institutions and NSOs is not limited only to these computations but covers a wide range of issues such as stakeholders' needs and to what extent the data is relevant to its intended purpose (such as to policy, assessment of development or anti-poverty programs, etc), how the basic concepts are defined (such as active employment or unemployed, clean water, etc), how often data and information are released within the determined time frame to be available to all stakeholders, and how stakeholders living anywhere and using a variety of dissemination techniques can access the best available and most recent data and information (for example to make decisions), how clear and understandable the available data and information are made to stakeholders (for

example by using meta-data and availability of contact persons actually engaged in generating the micro-data), etc.

The US Census Bureau and Loshin[132] define data quality as fitness for use. The Census Bureau uses three attributes to measure data quality: utility, objectivity, and integrity, which are defined across six dimensions: relevance, accuracy, timeliness, accessibility, interpretability, and transparency. Klugman[133] defines the quality of data used to make indicators for PRSP as consisting of "*…the whole range of factors that determine how well a particular indicator is suited to some use."*. These qualities, according to Klugman, include "*… how the indicator was collected, what it covers, how accurate or reliable it is, how often it is published, the time period to which it refers, and the level of aggregation.*[134] This in turn calls for *"…assessing how well the statistical system generates the data needed for PRSP indicators."*[135]

The IMF data quality assessment framework (DQAF)[136] and Carson on the other hand names five dimensions of data quality: integrity, methodological soundness, accuracy and reliability, serviceability, and accessibility, with elements and indicators as shown in the table found in Appendix C. EUROSTAT uses the following six dimensions of data quality namely relevance, accuracy, comparability, coherence, timeliness and punctuality, and accessibility and clarity[137]. Laliberté et al[138] compared the data quality approaches of the IMF and EUROSTAT for similarities and differences. They note that there are six common aspects to data quality namely relevance, geographical comparability/methodological soundness, accuracy and reliability, consistency/coherence, timeliness, and accessibility. Among their findings they indicate that the two approaches complement each other, with the IMF (with process-oriented indicators) focusing on qualitative aspects and EUROSTAT (with output-oriented indicators) focusing on providing quantitative measures. They also note that the

[132] Census Bureau 2006, pp. 1-2; Loshin 2001, p.10
[133] Klugman 2002, p.165
[134] Klugman 2002, p.165
[135] Ibid.
[136]Carson 2001, pp.20-23. Also See the IMF website for the July 2003 version of the DQAF which is a refinement of the preceding July 2001 version. http://dsbb.imf.org/Applications/web/dqrs/dqrsdqaf/ , retrieved August 26, 2009
[137] EUROSTAT 2009
[138]Laliberté et al 2004

IMF's DQAF takes a holistic view of data quality, whereas EUROSTAT focuses mainly on measuring the outcomes of selected statistical production processes (such as sample surveys, census, statistical processes using administrative sources, price index processes and statistical compilation). They further show that the Prerequisites of Quality and Assurance of Integrity dimensions of the IMF that focus on institutional and organizational aspects of data production and assurance integrity are not covered by the EUROSTAT approach.

As part of data quality, releasing survey results early by narrowing the time between data collection and publication enables stakeholders engaged in their respective fields to have a full picture of the extent of poverty at national, sub-regional, and global levels. A number of stakeholders however indicated problems they encountered regarding household surveys that can be related to the IMF's data quality dimensions. Chen and Ravallion[139] indicated for example that in order to assess poverty they have used primary data from about 675 surveys from 115 countries' government statistical offices. They indicated that the lag (in terms of frequency and access to surveys) of surveys has decreased globally (from a mean of 1.6 years in 2001 to 0.6 years in 2005), with the lag in the Middle East and North Africa due more to *slow* access to existing surveys rather than to the frequency of those surveys, while in Sub-Saharan Africa it was due more to the infrequent production of adequate surveys.[140] These figures relate to two of the IMF's DQAF dimensions, i.e., *accuracy* (reliability), and *accessibility*.

In the case of the CSA of Ethiopia, the HICES that was conducted in 2004/05 takes long to be published and made available to stakeholders for further analysis due to its complex nature and sensitivity[141]. As stated earlier, the HICES provides information on the income dimension of poverty. Another example is the agricultural sample enumeration survey (also known as agricultural census) conducted in 2000, which took even longer to be released. Improving delivery time is obviously essential when presenting poverty status information of countries all over the world to both local and global audiences.

[139] Chen & Ravallion 2008
[140] Chen & Ravallion 2008, p.18
[141] CSA 2007a, p.23

Poor quality data from data producers such as the national statistical offices and line ministries may stem from a number of factors such as the structure of the questionnaire, methodology, the presence of training manuals and how efficiently the training was given to enumerators, sound data collection, cooperation of respondents, cooperation of lower level administrative level officials, logistics, data entry, data cleaning and tabulation techniques, data storage and dissemination methods, and team spirit. In addition, financial and moral incentives to employees play important roles in maintaining data quality. Other factors may include availability and efficiency of human resources in the various departments, and availability and use of the right ICT infrastructure to support the statistical process of collection, processing, storage, and dissemination of data and information. As indicated by Klugman[142] the statistical system's human resources – the people who work for the component organization and the skills and expertise they posses – represent the most valuable and the scarcest resource. She also emphasized the important role that ICTs play in improving the efficiency of statistical agencies by providing opportunities for reducing delays in data processing, for dramatically reducing the cost of data dissemination through the use of technologies such as the Internet and CD-ROMs, and for expanding the scope for linking together different datasets.[143]

The summary below is compiled from reports from the different organizations which are external stakeholders to NSOs. The indicators prepared by these stakeholders using household and establishment survey data, and administrative records that are compiled from different sources including NSOs. One of the purposes of these reports is to inform the global audience, including countries themselves, about the socio-economic progress and point out the weak areas that require improvement.

In both its 2008 and 2009 Millennium Development Goals reports, the UN also stresses the importance of building a strong statistical system in developing countries to produce, analyze and disseminate data so that it is possible to measure, monitor and report on progress towards the MDGs[144]. The world development indicator (WDI) is a report released annually and is prepared by the WB. The data is collected from national statistical offices through household

[142] Klugman 2002, pp.172-173
[143] Klugman 2002, pp.172-173
[144] UN 2008, p.50; UN 2009, p.54

surveys and from central banks, customer services and line ministries as administrative records. According to the various reports of the WDI one data quality problem is *availability* of data, mainly in Sub-Saharan Africa (SSA), the Middle East and North Africa (MENA)[145]. The WB pointed out that availability is improving in general with time. In 2005 and 2006 24 out of 48 countries in SSA had at least one data set available – this figure dropped to 21 in 2008. In the MENA region, three countries had at least one data set available in 2007. The WB stressed that there were availability problems in consumption data in general. The WB also indicated that similar survey data from various countries was not comparable due to differences in timing, quality and training of enumerators. Other problems related to data quality include *selective non-response* and the length of period during which respondents must recall their expenditures in household income and expenditure surveys[146]. Some of the reasons for these problems were weak statistical systems, differences in statistical methods, coverage, practices and definitions.

[145] World Bank 2006, p.74, World Bank 2007, p.64
[146] WDI 2005; WDI 2006; WDI 2007; WDI 2008; WDI 2009

Indicators	Owner	Type of data	Data producers	Informational base (Specific survey or administrative record)	Reported weakness of data quality
Human Development Index	UNDP	Household surveys	NSOs	WMS, HICES	Availability
(HDI), Human Poverty Index (HPI-1)		Admin. records	Line ministries	National accounts	Availability
Information Communication Technologies index (ICT-I)	ITU	Household surveys	NSOs	WMS	Delay
		Administrative records	Telecommunication agencies and corporations, ITU project (in the case of Ethiopia ETA, ETC)	Reports	Delay
Millennium Development Goal (MDG)	UN	Household surveys	NSOs	DHS, Core Welfare Indicators Questionnaires (CWIQ), Household Budget Survey or HICES, Population Census	Availability[147] – either data is not produced at all or it is produced but not available to specialized agencies.
		Admin. records	Line ministries	Health, Education,	-
World Development Indicators (WDI)	World Bank	Household surveys	NSOs	HICES, population census, DHS, vital registration, agricultural census, trade and industry	The different editions of the WDI are not Comparable. Availability(data lags) , consistency, and reliability due to weak statistical systems in developing countries.
		Admin. records	Line ministries and sectors	Central banks, customs services	Availability -data lags

Table 0-1 Data quality problems by publication, data source, and informational base

(Source: author's interpretation summarized)

From the point of view of NSOs based on the DQAF, the problems indicated above could be categorized in either as accessibility (in terms of availability), serviceability (in terms of lag), or accuracy and reliability (in terms of unit or item non-response rates).

[147] See The Millennium Development Goals Report 2008, page 50.

4.3 Data, Information and Knowledge

Data, information, and knowledge are basic concepts used in the discussions of knowledge management both in theory and practice. In some literatures the terms data and information on the one hand and information and knowledge on the other are seen as interchangeable, but they are in fact distinct terms. Davlin[148] notes that while the words *information* and *knowledge* are often used interchangeably; they are not the same. *Information* does not mean the same as *data* either, though those two terms are often confused. Checkland and Holwell[149] indicate that the creation of information is done when humans attribute meaning to data. Davenport and Prusak[150] also stated that data, information, and knowledge are not interchangeable concepts. They perceive them as existing on a continuum, with information somewhere in the middle, linking data and knowledge; data, information, and knowledge are seen as not easy to distinguish in practice; at best one can construct a continuum of the three[151]. In a similar vein Max Boisot describes the distinction between data, information, and knowledge and how one may be transformed to another: "*….in the first step, noise is filtered out from incoming signals by the agent's sensory apparatus and the latter gets registered as data by different senses. In the second, non-information bearing data gets filtered out by the agent's conceptual apparatus, so that only information-bearing data is left to impact their action systems and thus get metabolized into knowledge."*[152]

Alavi and Leidner describe the difference between information and knowledge as follows; *"What is key to effectively distinguishing between information and knowledge is not found in the content, structure, accuracy, or utility of the supposed information or knowledge. Rather, knowledge is information possessed in the mind of individuals: it is personalized information*

[148] Davlin 1997, P.14
[149] Checkland & Holwell 1998, pp.91-93
[150] Nonaka I 2005, p.301
[151] Davenport & Prusak 1997, pp.8-9
[152] Rooney et al. 2005, p.179

(which may or may not be new, unique, useful, or accurate) related to facts, procedures, concepts, interpretations, ideas, observations, and judgements."[153]

This study takes the position of Davenport and Prusak and adopts their definitions of data, information, and knowledge[154] :

- Data is a set of discrete, objective facts about events.
- Data becomes Information when its creator adds meaning.
- Knowledge is a fluid mix of framed experience, values, contextual information, and expert insight that provides a framework for evaluating and incorporating new experiences and information. It originates and is applied in the minds of knowers. In organizations, it often becomes embedded not only in documents or repositories but also in organizational routines, processes, practices, and norms.

If the informational base used to track deprivation of capability in one or more of the dimensions of poverty such as the right to live long and healthy life, or be literate and numerate, etc has been determined, then the different types of data related to the dimensions can be generated by NSOs from household surveys that focus on individuals within a household like woman and child modules (regarding fertility and child mortality) from DHS, household surveys such as ownership to goods and access to services or household income from WMS and HICES, or community-based surveys such as issues related to migration, epidemic, social services from surveys administered for example to peasant association leaders in rural areas. Other related data can also be generated by line ministries in the form of administrative records with respect to health, education, ICT, etc. The collected data is then converted into information by categorizing the units of analysis (such as the woman, household, or the community leader), correcting the data by means of appropriate algorithms, and disaggregating it into more informative categories like gender, age, income, and administrative levels in order to inform policies for appropriate actions.

Data is contained in documents, various electronic storage media such as tapes and discs, or in information systems (in organizations, or ICTs). Davlin states that "… *data exists in paper*

[153] Nonaka I 2005, pp.165-166
[154] Nonaka I., vol. III, pp.302-305

and computer discs."[155] Boisot[156] also states that "… *data is borne on some physical substrate, whether it be stone, paper, or pure electromagnetic waves; information exists in the collective minds of a society and knowledge exists in an individual person's mind."*[157] In other words: "W*e obtain knowledge from individual or groups of knowers"*[158] Data is typically collected by NSOs from households and establishments and may include measures such as individuals' demographic characteristics, quantity of grains consumed or non-food items used, quantity of crops grown and the area covered by these crops, quantity of goods produced in an establishment or factory, or quantity of a certain food or drink produced and sold in an informal way, etc. These are the building blocks from which information is generated in a number of input-process-output cycles.

Although the use of ICT enables us to produce accurate data and information quickly, in practice it cannot replace the human mind when it comes to the analysis of empirical findings and subsequent decision-making. It takes humans to data or observed measurements (obtained for example from various surveys) or information into some meaningful context in order to perform some useful action. It is only with the involvement of individuals or data producing agencies that the transformation of data into information, and information into knowledge that can be communicated between different stakeholders becomes possible.

Boisot[159] defines players as family, organization, or society at large and goes on to explain that "… *intelligent players convert data into information and thence into knowledge through a two-step filtering process that is guided by the possession of prior knowledge." He continues by saying that* "*… ICT aims to facilitate the structuring of information and hence its flow."*[160]. Davlin also states that knowledge*"… is information possessed in a form that makes it available for immediate use…turning information into knowledge requires recognition of – and familiarity with – the relevant contexts and a mastery of the appropriate constraints. Some of these constraints are the ones that provide the link between information and its representation (that is, between information and data)…cultural contexts and*

[155] Davlin 1997, p.16
[156] Boisot, M. 1998
[157] Davlin 1997, p.17
[158] Nonaka I, vol. III 2005, p.305
[159] Boisot, M. 1998
[160] Rooney et al. 2005, p.179

psychological and social constraints also play a role[161]. Davenport and Prusak state that "… *as with any message, information has a sender and a receiver…a memo full of unconnected ramblings may be considered 'information' by the writer but judged to be noise by the recipient."*[162]

The value of information is perceived when its quality (integrity, methodological soundness, accuracy and reliability, serviceability and accessibility) has been preserved or improved. NSOs collect, process, store, and disseminate data and meaningful information. It is therefore important to understand whether data quality can be improved by studying the information systems used by NSOs in general and by using various ICTs within the information system in particular.

4.4 ICTs in Improving Data Quality

Information systems in organizations and the environment are defined as a set of interrelated components or elements that collect, process, and disseminate data and information[163]. Data and information about people, places, and things is collected, processed, stored and disseminated in order to facilitate planning, control, coordination, analysis, and decision-making[164]. Checkland and Holwell[165] note that the existence of Information Systems in organisations is to support individuals or groups of individuals in the organisation and the organisation itself. This is done in cycles of input-process-output with an evaluative feedback mechanism to correct the input stage and meet the desired objective as shown in the figure below[166]. Outputs from one subsystem may serve as inputs to another subsystem.

[161] Davline 1999, p.152
[162] Nonaka I, vol. III 2005, p.303
[163] Stair & Reynolds 2001, p.13; Laudon and Lanudon 1998, p.7
[164] Laudon & Lanudon 1998, p.7
[165] Checkland & Holwell 1998, pp.65,99
[166] Stair & Reynolds 2001, p.13; Laudon and Lanudon 1998, pp.7-11

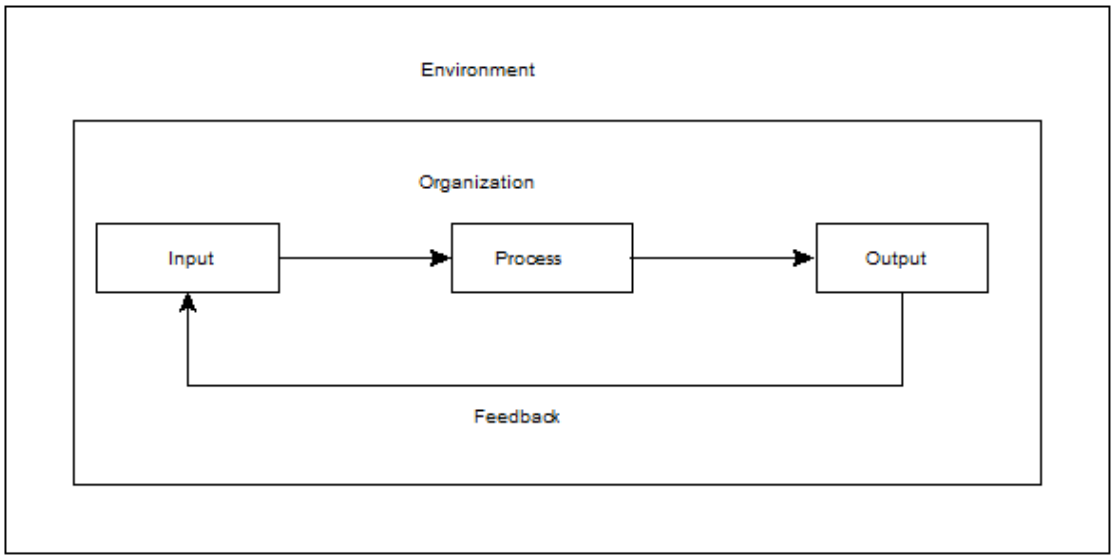


Figure 4-1 Components of an information system

(Source: Staire & Reynolds, 2001, Principles of Information Systems: a managerial approach)

The above diagram depicts the input-process-output cycle as discussed by Laudon and Laudon. During the input process a variety of intangibles (such as time, energy, skill) and tangibles (such as equipment, and data from the environment or an output from a different subsystem from within the organization) is fed into the system or subsystem; in the processing stage the captured raw data is converted into meaningful form by calculating, categorizing, comparing, storing, retrieving, correcting, and summarizing it manually or by using computers; the output process transfers the processed information to the people who use it in a presentable fashion (such as documents or reports), or it is used as feedback to the same subsystem and input to a different subsystem[167]. The production of data and information by NSOs passes through a number of input-process-output cycles with feedbacks at the various stages of the statistical process. Different inputs are fed into each process in order to produce different outputs. The output of one subsystem serves as input to the other subsystem, and this continues until the final product, the intended data and information, is produced.

A number of standard quality measures are taken in the different subsystems. Some of them are included in the process and inputs in the questionnaire (such as filtering questions or controls, ranges, codebooks, clear concepts and definitions, etc.), proper training, and

[167] Stair & Reynolds 2001, pp.8-14; Laudon and Lanudon 1998, p.8

frequent supervision, manual and computer edit checks, outlier detection and imputations, etc. Most decisions made by governments require not only the availability of data but also that the data should be *relevant, reliable and accurate.* The assessment of human development, socio-economic development and evaluation of policies based on any one of the informational base (such as from utilitarian, libertarian, or capability approaches) is also affected by the quality of poverty-related data collected. The quality of data collected by NSOs, mainly those of quantitative in nature, could be affected in any of the statistical stages or subsystems during pre-data collection, data collection, or post-data collection stages due to human factors (when questionnaires are designed, or trainers, or enumerators are trained) or non-human factors (faulty instruments, questionnaires, software or hardware). The inclusion or exclusion of questions, their wording, concepts and definitions of terminologies, and the flow and structure of the questions in household and establishment surveys conducted by NSOs are some of the data quality issues that need attention during the pre-data collection stages. Data quality problems during the data collection stage could be related to *accuracy* and *reliability* (such as coverage, sample error, response error, and non-sampling error). During the post-data-collection stage problems related to data processing (during data capturing, data cleaning, or tabulation), and *dissemination* could affect the quality of the data. Data and information that are disseminated through a variety of media reach wider audiences and may result in the production of more additional related information by external stakeholders, creating value as a result. The collection, retrieval, processing, storing, and dissemination of data and information bring together various ICTs, with the data to be collected, the activities, players with different backgrounds and skills (such as respondents, employees of data producing agencies, and international organizations, etc) interacting with one another.

Information systems in organizations using computers may consist of people, structures, strategies, policies, methods, rules, politics, and culture, and technology (such as computers, computer software, various storage and communication technologies)[168]. Technology is becoming a substantially important constituent part in information systems. The World Bank Group in its ICT strategy[169] defines ICT as a system that "…*consist[s] of hardware, software, networks, and media for collection, storage, processing, transmission, and presentation of*

[168] Laudon and Lanudon 1998, p.12-13; Stair & Reynolds 2001, pp.15-17
[169] World Bank 2002

information (voice, data, text, images)".It enables organizations involved in collecting information to collect data from different individuals and institutions on a variety of socio-economic and demographic topics through interviews and other methods, transcribing them into electronic formats, and processing and analysing them to produce information. Through the use of appropriate hardware and software technologies, organizations are also enabled to store and disseminate information for stakeholder consumption. The amount of data, information, and stocks of knowledge both at the global and national level are not only increasing over time, but stakeholders also demand improved quality. When this stock is managed well it enhances the decision-making process: quality data and information can be used to help achieve the goals set by a country's economic planning (e.g. in PRSPs). This requires good ICT infrastructure along with the various skills, systems and the expertise that can improve the collection, processing, storing, archiving and disseminating of data and information.

The use of appropriate ICTs in many spheres of activity at the individual and institutional level would be more productive if it helped the strategic goals of the organization to be met. For example, the MDGs have eight goals, 21 targets and 61 indicators. Eradicating extreme poverty and hunger is the first goal: it has a target of reducing by half the proportion of people who suffer from hunger with the proportion of the population below minimum level of dietary energy consumption as one of its indicators[170]. Governments in most developing countries are tracking this indicator, which obviously requires the collection and processing of relevant variables such as a basket of consumed food items by means of household surveys.

ICTs are used at different levels in the process of monitoring anti-poverty programs to assess the impacts of policies by measuring countries' progress towards meeting MDGs such as eradicating extreme poverty and starvation. This process involves a network of stakeholders all over the world: donors, both multilateral and bilateral, aid agencies and consultants working mostly in international organizations such as the WB, UN, and IMF, and various other experts and researchers working for governments and NGOs in different countries. The NSOs in many countries, as part of this process, are involved in the collection, processing,

[170] See http://www.undp.org/mdg/goal1.shtml

storing, and disseminating of socio-economic and other poverty-related data and information used to measure and evaluate the effective implementation and progress of various programs in the PRSPs and MDGs such as the proportion of people below the national poverty datum line. Some of the activities in this whole process in general and that of the NSOs in particular are difficult to perform without the use of appropriate ICTs. The processing of a huge set of data such as a population and housing census or a small survey that is complex in nature is not only difficult but also requires better and appropriate ICTs (such as computers, software, communication media), skilled personnel, and a well-established system to preserve data band improve its quality. Different resources are employed during the pre-data collection, data collection, and post-data collection stages to generate socio-economic, demographic and agricultural data in typical household and establishment surveys conducted by most NSOs as discussed below.

During Pre-Data Collection

A typical household survey conducted by NSOs passes through a number of input-process-output cycles within an open and complex system. The planning and organization of surveys involve different stakeholders from different walks of life, making a number of decisions regarding the overall operations of the survey, through consultations and meetings. This may include determining whether the statistical office has the authorisation to collect, process and disseminate the data, how relevant is the data that is going to be collected will be to the needs of stakeholders for monitoring and evaluation, and whether the required resources to conduct the survey are in place. Such decisions, among other things, may include budgeting for the survey (such as for vehicles, manpower, computing devices, salary, papers and printing, training, workshop, etc), the design of questionnaires and tabulation plans, preparation of training manuals and a complete list of code books for data elements, preparation of maps where sampled households live, development of the information system that will be used to capture, clean, and produce tables, the type of dissemination media, and the various quality control measures. In other words all the requirements for conducting the survey that may have an impact on the *prerequisites of quality* in the data quality dimension in the DQAF are addressed.

The design of questionnaires and tabulation plans, preparation of training manuals and codebooks, and training and pilot testing are some of the activities performed. The design of questionnaires and tabulation plans are usually the first activities that are done by the responsible survey team which may include the subject matter specialists, sampling and methodology experts, and systems analysts and programmers. It is worth noting that the quality of data can already be affected during the design of questionnaire and tabulation. To improve data quality teams hold frequent discussions to keep the concepts and definitions clear and simple, standardization of codes uniform, and ensures that the designs of the questionnaires and tables can be used in the computerised information system that will be used to capture the data and produce tables. The understanding of concepts and definitions such as households, enumeration areas, head of household, clean water, active employment, improved seed, plot size, etc in household surveys improve one of the second data quality dimensions, namely methodological *soundness* in the DQAF. The third data quality dimension, *accuracy and reliability*, with respect to non-sampling error, can be improved by reducing measurement and processing errors. Measurement and processing errors could be the result of poor questionnaire design such as lack of or wrong skip patterns, misleading controls for data capture, either through keyboard entry or Computer Assisted Personal Interviewing (CAPI).

Different ICTs are used in the pre-data collection activity. Hardware may include computers and laptops, printers, and uninterruptible power supplies (UPS). Different word-processing and spreadsheet software such as Microsoft Word and Excel could be used in designing questionnaires and training manuals. Communication (or networking technology) such as local area networks (LAN) maintain the continuous flow of notes and files among the survey team until the design is complete. Printouts of the questionnaire are made when the final version is ready. Some NSOs do not use paper for data collection; rather they design a questionnaire on their desktop computers and copy it onto laptops or handheld computers such as PDAs, which are later used to collect data by means of interviews or self administered in Paper and Pencil Interviews (PAPIs), CAPIs, or in Web-based mode.

During Data Collection

Household survey data can be collected in different ways, generally by conducting interviews or even self-administered, where a household member fills in a questionnaire. Some of the methods of data collection used by different countries include Paper and Pencil Interview (PAPI), Computer Assisted Personal Interview (CAPI), Computer Assisted Telephone Interview (CATI), Computer Assisted Self Interview (CASI), mails and delivery, and by Internet[171].

In a face-to-face interview, PAPI by far is the most common method for collecting socio-economic, demographic and agricultural data and has been employed for a long time especially by most NSOs in Sub-Saharan Africa. In this mode, data collectors (also known as enumerators) interview the respondent (an external stakeholder to an NSO) and fill in the questionnaires with a pencil with the aid of training manuals, codebooks, measuring instruments (compass, tape measure, scale, etc). The completed questionnaires are collected from each point and transported to the central office for keying in data and further processing. Face-to-face interviews can also be conducted using the CAPI or paperless mode, i.e. with the help of portable computers such as tablet PCs, laptops, handheld computers (PDA or its newest UMPC version), or mobile phones. The same questionnaire forms are loaded into the computer and the enumerator conducts the interview by reading the questions from the screen and entering the response using the stylus and virtual keyboard on the PDA or UMPC screen. This method eliminates errors while the enumerator is still in the field. Some studies indicate that it improves the timeliness, accuracy and reliability of data by reducing unit and item non-response rates. This can be done by using computer editing programs that capture and correct the data in the field. The routing of questions is maintained by skips allowing the enumerator to move to the right question number, page, or section. It has a long list of drop-down codes that can be called up more quickly than the PAPI mode. The CAPI mode involves a substantial initial cost for the portable computers, accessories, software, and time for training, but is economical in the long run by saving the cost of paper, printing and data entry, especially in quarterly, biannual, and annual surveys. In this mode, all data entry and some computer errors are corrected in the field using the portable computers. The digital data is

[171] Roberts outlined commonly used methods of collecting data from sampled units. See http://eprints.ncrm.ac.uk/418/1/MethodsReviewPaperNCRM-008.pdf

transmitted for further processing from the portable and other computers using unshielded twisted-pair (UTP), wireless technology and a modem[172], or by using cellular transmission over a telecommunication service. Alternatively data can also be stored on storage devices such as magnetic and optical disks, flash drives, etc and then copied onto a computer in the office. Computer Assisted Telephone Interviews (CATI), use existing telephone infrastructure for non-face-to-face interview in which the respondent is interviewed over the phone and the responses either recorded on tapes or documents and then digitised later or directly keyed into a desktop computer.

Self-administered PAPI and CAPI modes of data collection can also be used to collect data. The concerned household member (the head, a woman if the questionnaire is for women only, etc) is provided with the survey instrument for self-completion, which is collected later. Another self-administered mode of data collection is through mail or delivery of questionnaires to be filled in by the respondent to be collected later on. Web-based data collection, although not common in household surveys in developing countries, is also another way to collect data from respondents. Web-based data collection is growing in popularity in developed countries with the growing penetration rate of the Internet, and availability and decline in the cost of improved and user-friendly ICTs. This is a type of computer assisted self-administered mode of data collection where the respondent accesses the questionnaires, designed using various mark-up languages for presentation of the questionnaire such as hyper text mark-up language (HTML), extensible mark-up language (XML)[173] or others, using Internet browsers to fill in the forms and send the electronic copy to the data producing agency, either directly to the data base or through e-mails. A CAPI mode for both interview and self-administered methods can be developed using VISUAL BASIC, Blaise, CSPro, ASP.NET, Microsoft.NET, PHP, PERL software, etc. Some data producing agencies, mainly NSOs in developed countries such as the Netherlands, Belgium, Germany and Finland[174] use CAPI mode for laptop and Web-based data collection in household and establishment surveys. It is worth noting that these countries all have well-developed infrastructures in ICTs (both institutional and national) as well as in roads and

[172] A modem is a telecommunication device that converts digital signal (such as data entered in a portable computer) into an analogue signal (to be transported over a telecommunication line) and back to digital.
[173] Kunzler, Uwe 2002
[174] See for example: Bjorkqvist, Sven et.al 2002; OECD 2005, pp.19-20; OECD 2006; Roos, Marko 2002

electricity. The ownership and usage of PCs, Internet (with different bandwidth), mobile phones, pocket PCs and standard application software by households and their members are also increasing making it easier for NSOs in developed countries to use the CAPI mode and web-based data collection. The transition from PAPI to CAPI and its implementation using laptops or pocket PCs for the collection, editing, and transmission of data in these countries therefore would not be as difficult as countries in Sub-Saharan Africa.

A number of studies were conducted, mostly in developed countries, to see whether the use of ICTs (mainly using laptops and PDAs) for data collection in different types of surveys improve specific data quality dimensions such as accuracy and reliability, serviceability, etc) and save costs. Some of the studies compared the CAPI and PAPI modes in health information systems, agricultural and household surveys (for example Neas et al 2006; Schräpler et al 2006; Galliher et al, 2008; Vivoda and Eby 2006; Jodi et al. 2007). Some of these studies indicated that the CAPI mode of data collection was found to be better over the PAPI mode with respect to accuracy and reliability (focusing on unit and item non-response rates) while others maintain that the use of mobile data collection has little effect on data quality compared to the PAPI mode. Other organizations that produce data use the CAPI mode using laptops or PDAs to collect household data, despite all the problems due to these ICT's (for example MoH 2008).

Jodi et al applied the CAPI mode using PDAs in an insecticide-treated bed net (ITN) household survey in Togo (in 2005) and Niger (in 2006), conducted to evaluate ownership and usage by pregnant women and children under five[175]. The PDAs use Windows Mobile_ 2003 or Windows Mobile_ 5 operating systems, compact flash (CF) and secure data (SD) card slots, 10 hours' battery life with backup, an alternative power source (such as a battery pack, car charger, solar power), and are equipped with software applications such as GPS developed for navigation and mapping using VISUAL BASIC Visual Studio .NET 2005 (Microsoft Corporation) and an electronic questionnaire designed using Visual CE_9.0 Professional (Spyware Inc., Cambridge, MA) both developed at the Centre for Disease Control (CDC), The Microsoft Office Access 2003 database is used to store the GPS records and interviews, SAS 9.1 for processing and presentation of preliminary report, and Arc View

[175] Jodi et al 2007

3.3 (ESRI - Environmental Systems Research Institute, Redlands, CA) to monitor GPS data quality[176]. They noted that in Togo it took only 19 days to map 21,588 households and conduct 3,523 interviews in 162 EAs; in Niger it took 26 days in the field with 15-16 actual working days to map 28,552 households and conduct 1,801 interviews in 112 EAs with mean interview times and median distances travelled in the EA as follows[177]:

	Togo		**Niger**	
	Mapping	Interview	Mapping	Interview
Total days	19		15-16 working days	
Mean time per day per household	1 hour and 48 minutes (with a median of 40 maps by each enumerator)	7.5 minutes (each enumerator conducting 8–9 interviews)	3 hours 20 minutes (with 58 maps by each enumerator)	14 minutes (each enumerator conducting approximately 4 interviews)
Median distance travelled by enumerator in one day within the EA	1.3 km		3.3 km	

Table 0-2 Mean time and median distance travelled per day by interviewers for mapping and interview

(Source: compiled by the author)

They indicated that the combined use of GPS and CAPI applications on the PDAs enabled enumerators to include sampled households through the use of the GPS, and improve data quality by reducing errors through prompt feedback and resolving various data conflict problems during data entry in the field.

A similar CAPI application was used by the Ethiopian Federal Ministry of Health (MOH) when a malaria indicator household survey was conducted in 2007 (7, 621 households in 341 EAs) in collaboration with CSA, WHO, UNICEF, the centre for disease control (CDC) of USA, USAID, etc in order "… *to evaluate the progress of the national malaria control program with respect to ownership and application of insecticide-treated nets and anti-*

[176] Ibid., p.394
[177] Ibid., pp.395-396

malarial medicines, prevalence of fever, malaria parasitmia, and anaemia[178]. The CAPI mode of data collection was used and the design of the questionnaire and editing programs was done by the CDC using VISUAL BASIC and transferred to 113 PDAs (HP IPAQ HX249X and Dell Axim–51 with Windows Mobile 5.0 operating system). The PDAs were equipped with GPS receivers, which assist interviewers in listing households, compiling a random selection of sample 25 households from the listed households for interview within the EA, and even navigation to the selected household[179]. A 10-day training period about the survey and PDA/GPS-based data collection was given to the survey team (interviewers, team leaders, and supervisors) and 25 supervisors, 100 interviewers, and 25 drivers were deployed for data collection, whereas the supervisors inspect recorded questionnaires in the PDAs, inspect surveyed households by the enumerators using the GPS navigation system and do some minor maintenance such as restoring memory[180]. Some of the challenges encountered in the survey include insufficient time for logistic, human resources, hardware, and software problems, where some of the problems were solved in the field during data collection but could have been solved if sufficient time were given[181].

[178] MoH 2008
[179] Ibid., p.6
[180] Ibid., pp.8,54
[181] Ibid., pp.51-54

Category	Type of problem
Logistics	○ Delay in the supply of PDAs during the training time
Human resources	○ Small numbers of experts in programming and using the PDA
Hardware	○ Five out of 120 PDAs were unable to function ○ Village names were misspelled due to defective electronic key boards (and therefore they were forced to use the built-in key board) ○ Problems related to charging PDAs due to lack of road and electricity infrastructure (so forced to walk or ride on animal taking up to 5 hours to collect data) ○ Waiting to acquire signals by PDA GPS units ○ Partial compatibility of the CAPI application with Windows Vista operating system ○ PDA memory run-out during data collection while swapping the secure digital (SD) cards to merge data sets ○ Erroneous system date such as May 2005. which has a problem in calculating age
Software	○ The application needs a date in the questionnaire, but was not designed to accept the Ethiopian calendar variables. This has led to the use of a Gregorian to Ethiopian calendar converter program. ○ Uploading of chronology of historical events was not done on PDA due to time constraints ○ No GPS data (latitude, longitude, and altitude) for most houses although surveyors were navigating. ○ Persistence of programming errors such as skip patterns even during the actual data collection. ○ The CAPI application in the PDA did not automatically generate a unique ID number for every person who volunteered to provide a blood specimen in order to minimize errors arising from the manual coding of slides by technicians; as a result a hard copy was used to give ID manually.
Security	○ Inadvertent locking of passwords of enumerators, and also unintentional changing of passwords of PDAs by non-enumerators that were supposed to be managed by PDA administrators.

Table 0-3 Software and hardware problems encountered during CAPI mode of data collection using PDAs

(Source: summarized from Ministry of Health (MoH), 2008, Ethiopian malaria indicator survey 2007)

The US National Agricultural Statistics Service in 2003 used the CAPI mode for PDAs on the monthly cotton objective yield (COY) survey. Neas et al[182] investigated this mode to see if it improved data collection and data quality among other things. In addition to cost savings the reported benefits obtained included reduction of data collection error (an element in the *accuracy and reliability* data quality dimension) as a result of the use of automated edits on the field, quicker preliminary analysis (an element in the *serviceability* data quality dimension), and reusability of the data collection instrument for the next survey[183].

Schräpler et al examined the implication of the CAPI mode of data collection using laptops, on the *accuracy and reliability* dimension (focusing on unit non-response, and item non-response such as net income) and compared it to the PAPI mode, using the German Socio-Economic Panel longitudinal survey[184]. One of their findings indicates that the rate of implausible values (values wrongly encoded by interviewers) is lower in CAPI than in PAPI in general[185]. However, their comparison with respect to sensitive items such as gross income, results in higher item non-response rates (constituting *don't knows* by respondents who earn low salaries and refusals by respondents in high earning positions) in CAPI than in PAPI in general[186].

Galliher et al used both the paper form and PDAs (Sony Clie PEG-T615C with Palm operating system version 4.1); they used Pendragon Forms 3.2 software to develop the CAPI application form on the PDA, and compared the response rates and error rates in both cases in an office based collection of *immunization* data[187]. They found that in terms of completeness of data of the returned forms from each mode, PDAs produced more complete data (with 0.4% missing items) than the paper method (with 3.5% missing items), although the overall return rate (i.e. unit response rates) for PAPI mode (94%) is better than with the CAPI mode using PDAs (82%) but with more errors of omission (i.e. no response) in the PAPI (35%) mode than the CAPI mode using PDAs (3% observed).[188]

[182] Neas et al 2006
[183] Neas et al 2006, p.10
[184] Schräpler et al 2006
[185] Ibid., p.16
[186] Ibid., p.23
[187] Galliher et al. 2008
[188] Ibid., pp.155-157

Vivoda and Eby compared accuracy and speed of data collection using the PAPI and CAPI using PDAs (Palm Tungsten, Palm operating system, built-in capabilities of a cellular phone, wireless Internet, a small keyboard) modes by observing along roadways. They found out that both the CAPI and PAPI mode are comparable with respect to the accuracy and speed of data collection, but the use of PDAs is superior to the paper method in the rainy season. Therefore they recommended the use of PDAs for this type of survey. Their result shows PDAs can alternatively be used in data collection but also cautions that PDAs are more appropriate (appropriate *computing device* for the survey indicator in the pre requisite to data quality dimension) for *safety belt surveys.*

The use of CAPI mode using PDAs over the PAPI mode for data collection in different types of surveys produces a mixed result in terms of improving some of the data quality dimensions as seen above. Along with the use of PDAs or other portable PCs for data collection go the limitations such as battery life, visibility, loss of PDA, computer and technical literacy, troubleshooting in case of failure, power supply, and infrastructure. The transition from PAPI to CAPI needs careful planning and preparation in the selection of hardware and software including system software, and proper training in the usage and troubleshooting of the PDAs. In addition, sufficient time is necessary for intensive training so that enumerators are comfortable to use the PDAs or other mobile computers and apply their expertise in using the PAPI mode in the CAPI application. However it is also important to note that when actual data collection is implemented, especially in new surveys, some difficulties (different from the ones found during pilot testing) are likely to be encountered. The same type of situation can also be encountered in CAPI mode in addition to hardware and software problems. Availability of human resources to correct any hardware or software problems (programming, system problem) encountered and backup such as accessories including PDAs is also needed during the actual data collection to maintain the smooth functioning of the entire survey process.

The following table (table 4-4) gives a summarised interpretation of a comparison of CAPI and PAPI from the above literatures with respect to some elements such as *item* and *unit non-response* (completeness of data items and completeness of form) in the *accuracy and reliability* and other data quality dimensions in the DQAF.

Researchers	Survey type	Portable computers used	Effects of CAPI and PAPI modes	Data quality dimension CAPI improves
Galliher et al	Immunology	PDA	CAPI is better in improving item non-response. PAPI results in more unit responses but with more omissions of items	*Accuracy and reliability* with respect to improving item non-response
Vivoda and Eby	Safety belt	PDA	Accuracy and speed are comparable in both CAPI and PAPI modes	*Pr requisite to data quality dimension* with respect to commensurability of computing devices
Schräpler et al	Socio-Economic Panel longitudinal survey (household survey)	Laptop	PAPI is better for sensitive items such as gross and net income. CAPI is better in terms of improving incorrect values that may result from interviewer or instrument.	*Accuracy and reliability* except sensitive items.
Neas et al	Cotton Objective Yield Survey (COY)	PDA	CAPI is better as result of reduction of errors.	*Accuracy and reliability*

Table 0-4 Effects of CAPI and PAPI modes on data quality

(Source: compiled and interpreted by the author)

During Post-Data Collection

Depending on the mode of data collection, basic activities in the post data collection stage may include documentation, data editing and coding, data capturing, data cleaning, data preparation, tabulation (estimation of variables and errors), data basing, uploading files on the website, dissemination of survey results using publications, CDs, website. NSOs who use CAPI and web-based mode of data collection receive the collected and partially or fully cleaned data from the field over communication and networking technologies WAN, VPN, dial-up, satellite or cellular transmission on a real-mode or batch mode. NSOs may have DBMS in the central place where an online or offline data entry is done in the field. Others receive data from the field after data collection (entry and editing) is completed. The collected data can be sent by e-mail to the NSOs over a telephone line (using a modem), cellular, satellite, etc depending on the functionalities of the portable PC of NSOs ICT infrastructure and that of the national telecommunication infrastructure. In cases where there is such limitation on the infrastructure for connection and secure transmission of data, the collected data from fields can be delivered to the central office by copying the data onto storage media (such as diskettes, cards, memory sticks), or by returning the portable PC.

The remaining post data collection activities in the central offices involve further processing of the collected data such as further cleaning if needed and tabulation. ICTs used include laptops, PCs, servers with accompanying server and desktop operating systems, and word processors and spread sheets (for example Microsoft Word and Excel), other specialized software for data manipulation and tabulation (for example CSPro, Blaise, SPSS, SAS), database software (MS SQL, DB IV, MY SQL) to maintain, disseminate and for on-line access of survey data. Most commonly used specialized statistical packages such as SPSS, SAS, STATA, STATGEN can also be used for further analysis of data, etc.

The usefulness or the quality of data and information increases if they are communicated to various internal and external stakeholders. Statistical offices in developing countries are rich in socio-economic, demographic, and agricultural data and information which they have collected over a long period of time. ICTs in the post-data collection stage play a big role by improving the quality of data and information with respect to one of the data quality dimension *accessibility.* This can be done by improving the following indicators of the

accessibility data quality dimension, the clarity of the survey report, availability and adequacy of dissemination media, availability of meta-data, and availability of contact persons who know how the data is generated. The availability of instruments that facilitate the exchange of data and information is of greater benefit mainly to external stakeholders who live in geographically distant places such as international and continental organizations (the WB, and IMF, UN agencies) and other stakeholders in the higher and research institutions, professional and civic associations, researchers, respondents, and students in order to make further analysis and compare global trends with respect to poverty and other indicators.

The following table (Table 4-5) gives a summary interpretation of how various ICTs are used during the pre-data collection, data collection, and post-data collection stages in household surveys conducted by NSOs in general. The various information systems components (inputs, process, and outputs) used and the possible data quality dimensions that can be improved in the above stages are listed.

IS components and data quality	Statistical Processes		
	Pre-data collection	Data collection	Post-data collection
Input	Internal and external stakeholders	Enumerators and supervisors	Subject matter specialists, programmers, website and database managers, programmers, system analysts, data entry operators
Process	Discussion, design of questionnaire and tabulation (table formats), training of enumerators.	Data collection, supervision and transmission which can be done either using PAPI or through the use of ICTs.	Receiving data, sending online error messages during online data entry, Data capturing, data editing and tabulation. Archiving , data basing, and dissemination of various survey documents
Output	Questionnaire, CAPI application (on laptops, PDAs, web	Filled questionnaires, electronic data (on laptops, PDAs, web	Cleaned electronic data, statistical outputs, publications

	pages)	pages)	(on paper, CDs, websites)
Feedback	Feedback from discussion and Pilot testing	Supervision, computer assisted error messages.	Various feedbacks obtained in the form error messages during the process of data capturing using computing devices, editing, and tabulation, and also through discussion with internal and external stakeholders.
Data quality dimension	o Prerequisites of quality o Methodological soundness o Accuracy and reliability	o Accuracy and reliability o Serviceability	o Accuracy and reliability o Serviceability o Accessibility
ICTs used	*Hardware*: PCs, laptops, UPS, printers. *Software*: Operating systems, word processors and spreadsheets, system development tools for CAPI *Communication*: LAN	*Hardware*: Laptops, PDAs, mobile phones, servers. *Software*: Operating systems, word processors, DBMS (for web based remote database or direct data entry), spreadsheets, system development software such as Microsoft Visual Studio, ASP.NET, PHP, XML, HTML, VB. Software tools to develop data entry (for CAPI or direct entry) and cleaning and tabulation systems such as CSPro, Blaise, Pendragon, EPInfo, etc.	*Hardware*: Laptops, PCs, Servers, *Software*: Operating systems, word processors, DBMS (for web based remote database or direct data entry), spreadsheets, system development software such as Microsoft Visual Studio, ASP.NET, PHP, XML, HTML, VB. Software tools to develop data entry (for CAPI or direct entry) and cleaning and tabulation systems such as CSPro, Blaise, Pendragon, EPInfo, etc. Statistical software

		Communication: WAN, VPN[189], dial-up, satellite or cellular transmission of data over a telecommunication infrastructure for electronic data capture and transmission.	for data manipulation and analysis such as SPSS, STATA, SAS, etc. MS SQL, DB IV, MY SQL, etc to develop and maintain data bases, and disseminate the outputs. *Communication*: WAN, VPN[190], dial-up, satellite or cellular transmission of data over a telecommunication infrastructure for electronic data capture and transmission. LAN, routers, bridges, hubs, network cables for file sharing in the office.

Table 0-5 IS components and data quality in the different stages of the statistical process.

(Source: summarized by the author)

4.5 Conclusion

In this chapter different approaches to data quality were discussed and the IMF's DQAF is adopted for this study. The different types of ICTs that are used in the collection, processing, storing, archiving, analyzing, and dissemination of data and information were also discussed, also in relation to improving the different data quality dimensions. Some of these dimensions include accuracy and reliability, accessibility, and serviceability. The quality of data and information produced by NSOs can be improved with the use of an appropriate information system, such as one that contains appropriate and better ICT infrastructure and human capital.

[189] Virtual private network (VPN) is a secure connection across the Internet, by outsourcing one's WAN to save money and equipment to Internet service providers (ISPs), to transfer encapsulated information between receivers and senders.

[190] Virtual private network (VPN) is a secure connection across the Internet, by outsourcing one's WAN to save money and equipment to Internet service providers (ISPs), to transfer encapsulated information between receivers and senders.

Field data collection using portable PCs (laptops, PDAs), web-based or stand-alone CAPI mode, was also discussed. Some NSOs in developed countries have used the full range of functionalities of portable PCS over well-developed national telecommunication, power, and institutional ICT infrastructures and realized improvements in data quality. Although the use of mobile data collection in household survey data collection by NSOs and other data producing agencies in developing countries, is still a debatable issue raising mixed feelings, some studies show that significant improvements were made in specific data quality dimensions, mainly in reducing errors resulting from non-response, a part of the *accuracy and reliability* data quality dimension. They further indicated that in some surveys, the time spent during data collection was reduced, although not substantially, by avoiding separate data entry and a varying degree of data cleaning as done concurrently during data collection. Under the PAPI mode the detection and correction of erroneous data is done in the head office during the post-data collection process while the data is manually edited or cleaned using computer programs, long after data collection has been done. This makes it difficult to check errors on the spot. However the transition from PAPI interview to CAPI interview needs careful and comprehensive planning and assessment by NSOs. For this reason a description of the CSA of Ethiopia, focusing on the roles in PRSPs, its objectives, strategies, organizational structure, types of surveys conducted, ICT infrastructures, etc. is discussed in the following chapter. It is followed by another chapter that describes further details on how data is generated and the flow of data and information, and the interactions between the different stakeholders within CSA.

CHAPTER FIVE

CENTRAL STATISTICAL AGENCY OF ETHIOPIA AND ITS ACTIVITIES

5.1 Introduction

In this chapter, the roles of CSA and its relationship with the main stakeholders from the point of view of data quality are discussed. The chapter also deals with the contribution of the CSA of Ethiopia in the poverty reduction strategy paper (PRSP), followed by a discussion of ICT strategies of the Government as covered in the second phase of the PRSP of Ethiopia (also known as plan for accelerated and sustained development to end poverty or PASDEP). The CSA's ICT infrastructure and the types and usage of various ICTs at different phases in the statistical process during pre-data collection, data collection, and post-data collection (processing, storage, and dissemination) stages are discussed next, followed by the conclusions drawn from the chapter.

CSA produces poverty related data and information to be used by different stakeholders such as ministries, donors, international institutions such as the WB and IMF, NGOs, etc in designing and implementing various interventions in education, health, food security, etc. Estimates of the proportion of the poor who fall below a certain poverty line are measured using data collected, processed, and disseminated by the CSA. Periodic surveys such as the household income consumption and expenditure survey (HICES) and the welfare monitoring survey (WMS) are used to provide socio-economic indicators. The WMS for example provides the non-income dimension of poverty and is used as input to the process of monitoring and evaluating PRSPs. The income dimension of poverty is obtained from the HICES and is used to compute total national poverty, taking in to account both the food and non-food requirements. Different departments in the CSA are responsible for producing data and information. The following section discusses the role of CSA and the responsibilities of the different departments.

5.2 Role of CSA and Stakeholders

The CSA operates under the auspices of the Ministry of Finance and Economic Development (MOFED). It was re-established according to the Proclamation endorsed by the House of People's Representatives on April 2005[191] with the following objectives:

- To collect, process, analyze and disseminate the necessary socio-economic and demographic statistical data through censuses, sample surveys, continuous registration and administrative recording systems; and
- To provide technical guidance and assistance to government agencies and institutions in their endeavour to establish administrative recording, registration and reporting systems; and build the capacity required for providing directives and consultations in database creation and development of administrative records and registration systems.

The CSA is headed by the director-general, assisted by three deputy directors-general responsible for economic statistics, demographic and social statistics, and operations, methodology and data processing. The different socio-economic and demographic surveys undertaken by CSA are listed below under their respective departments.

Economic statistics

- Natural Resource and Agricultural Statistics department

  The annual surveys are on crop production forecast, crop production for the long rainy season, land utilization, farm management, livestock, poultry and beehives survey, crop production in the short rainy season, and large and medium scale commercial farms. The agricultural sample enumeration survey, which covers all the above annual agricultural surveys in addition to other components, is conducted every ten years.

- Industry, Trade and Service Statistics department

  A manufactured product producers' price survey is conducted quarterly and the medium and large scale manufacturing industries survey annually. Distributive trade

[191] CSA 2005a, p.3. Can also be accessed from http://www.csa.gov.et/Policy.htm

and services, mining and quarrying, small-scale manufacturing industries, and cottage/handicraft manufacturing industries surveys are conducted every three years. A census of economic establishments/enterprises is done every five years.

- Household Budget, Welfare, Manufacturing and Price Statistics department.

  Agricultural commodities producer prices and retail prices of goods and services are surveyed every month. A household income, consumption and expenditure survey is conducted every five years, and a welfare monitoring survey every three years. In addition, the core welfare monitoring questionnaire survey is conducted annually.

Demographic and Social Statistics:

- Population and Housing Statistics department,

  Demographic and health surveys are conducted every five years, and a population and housing census is conducted every ten years with the participation of all staff of the CSA and the Population Commission.

- Vital Registration and Statistics Department, and
- Vital Events Statistics.
- Manpower and Social Statistics department.

  A current employment survey is conducted every six months and a labour force survey every five years.

Operation, Methodology and Data Processing:

- Field Operation department,
- Data Processing department,
- ICT Development department, and
- Printing Services.

In addition to the above surveys, concerned departments also prepare statistical reports, price indices and analytical reports. The field operation, data processing and methodology

departments and printing services are involved in all surveys, providing support through logistics, data processing activities, sampling, and publishing survey results.

The CSA, with the help of inputs and needs from stakeholders, is responsible for designing questionnaires, collecting socio-economic and demographic data from the field, and editing, cleaning and processing it to produce information for dissemination using different media. The MOFED on the other hand oversees the implementation of different survey activities performed by CSA, use the information and indicators from CSA and other line ministries as inputs to make further analysis, plan economic policies, and use them to monitor and evaluate economic development.

In addition to collecting socio-economic data, the CSA delivers informative services, both at the head office and through its website. The various services include information on schedules of future surveys which enables stakeholders who are interested in the specific survey results to budget their own resources. Hard copies of survey publications, administrative records, and statistical abstracts are available at the head office for sale whereas their electronic versions are downloadable from the website. Provision of raw data is also done for payment based on an access policy for raw data approved by the council of ministers[192].

The CSA classes stakeholders as internal and external, and describes them as individuals, groups, or organizations who have an interest and concern in statistical information[193]. It can also allocate budgets and provide technical and material assistance. In addition, the CSA collaborates with them while collecting socio-economic used to formulate, monitor and evaluate policies and programs. For example, MoFED needs a continuous flow of timely, accurate and reliable data and information on national accounts and socio-economic indicators to be used in economic planning and monitoring. MoFED may not be able to follow up economic growth without reliable or complete information, and may review the organizational structure of the CSA and its budget as a result. International organizations also resort to the use of other data sources if the information from CSA is not reliable, and may

[192] CSA 2004a
[193] CSA 2005b, pp.21-22 (available only in the local language)

reduce assistance as a result. More importantly, respondents may not cooperate in the future if the information they provided is misused or was not kept confidential. This means that the CSA needs to perform more comprehensive data quality improvements to satisfy the needs of stakeholders. Some of the stakeholders for CSA are listed below.[194]

Internal:

- The director, deputy directors, department and service heads, and
- Staff both at the head office and branch offices.

External:

- MoFED, government offices and institutions – including most line ministries, authorities, commissions and agencies.
- Regional governments (nine), Addis Ababa regional council, Dire Dawa regional council, and various offices established under them
- International and continental organizations
- NGOs
- Higher institutions [e.g. universities], research institutions [e.g. federal and regional agricultural research institutions]
- Government and private producers and service providers
- Respondents
- Professional, public, and civic associations and
- Researchers.

Different stakeholders have different needs. They may want more items or variables in specific surveys or censuses, which they believe to be vital to their operation. Some variables in the draft questionnaire designed with inputs from external stakeholders, for example from the WB, might be excluded in order to align it with the economic plan of the country. Although it is not always possible to satisfy the needs of stakeholders, data quality must be maintained regardless of the type of survey or census.

[194]Ibid., pp.21-22

Loshin states that organizations who own the data are accountable for maintaining its quality[195]. The CSA, as a government agency, is the owner of the data and information that it generates through its annual and periodic household and establishment surveys and censuses. The CSA renders different services to stakeholders that include providing advisory services on statistical activities; providing assistance in guiding and coordinating the statistical work pertaining to the collection, organization, analysis, and preparation of publication and dissemination of statistical data, making stakeholders cover certain costs and setting priorities, undertaking specific surveys falling outside the scope of its plans and programs[196]. The data the CSA produces on behalf of one of the external stakeholders mentioned in the above list is, however, owned by the stakeholders. The quality of the data they produce is therefore maintained by them before the CSA receives it in the form of administrative records. It is noted in the five-year Ethiopian national statistical development strategy (NSDS) document that over 40 organizations were identified in the National Statistical System (NSS) as producing statistical data and publications[197]. The second strategic theme in the NSDS calls for the establishment of a unit in the CSA that coordinates the NSS and would be responsible to develop a country-specific data quality assessment framework applicable to members of the NSS[198].

5.3 Role of the CSA in PRSP

Data collection, processing, storage, and dissemination are the main activities of most national statistical offices. Loshin states that "… *strategic decisions based on untrustworthy information are likely to result in poor decisions."*[199]The CSA conducts a number of socio-economic surveys and the formulation of the first and second phases of the PRSP, called SDPRP and PASDEP respectively, was based on the evidence obtained from researches conducted on various socio-economic household surveys by the CSA and other administrative records compiled from line ministries.

[195] Loshin 2001, p.46
[196] CSA 2004a, p.4
[197]CSA 2009, pp.30, 41,114
[198]CSA 2009, pp.80-81
[199] Loshin 2001, p.10

The quality of data, as an important factor in the PRSPs produced by different ministries and national statistical offices, is not expected to be hundred percent perfect due to a number of factors. In the case of the CSA, the quality of data has improved over the years. In addition the time between data collection and publication of information data has been reduced. For example, the CSA has shown progress in the *timeliness* data quality dimension in the WMS surveys that were conducted in 1996, 1998, 2000, and 2004. The WMS and HICES along with others such as population and housing censuses, urban employment-unemployment surveys, and demographic and health surveys, are useful in measuring the poverty indices and profiles. As was seen in the previous chapter, the WB uses some of the data and information obtained from the NSOs and other line ministries to assess income poverty in order to collaborate with countries in the development of strategies and make decisions. Similarly, the UNDP also uses the data sets from NSOs and other line ministries to report on the multidimensional nature of poverty in its HDR.

Similarly, the MoFED uses quantitative and qualitative data obtained from CSA and other line ministries to monitor and evaluate the various programs set out in the PASDEP. Broadly speaking, income poverty, as discussed in the previous chapter, contains the income and non-income dimensions, with generally two possible sources of information for these dimensions. A number of anti-poverty programmes aimed at improving income are implemented by various government line ministries. Monitoring the progress of these programs can be done by conducting household surveys that measure these dimensions. Such surveys may include the Household Income Consumption Expenditure (HICE), the Welfare Monitoring (WMS), and Participatory Poverty Assessment (PPA) surveys. These can act as feedback mechanisms to measure to what extent these strategies were implemented and indicate whether updated or new strategies are needed. NSOs as the source of nationally representative socio-economic data can, in collaboration with other stakeholders, generate more poverty-oriented data and information, which are used in the monitoring and evaluation of various programs in the PRSPs. The data and information produced are also used as inputs to assess poverty trends and planning future RRSPs. This means that the quality of data obtained from Statistical offices and line ministries that are used to make poverty-related indicators should be as good as possible. Assessing the different stages of the statistical process in which data and information are generated is therefore one important aspect in order to preserve and improve the quality of data and information.

5.4 ICT Strategies of the Government and CSA

5.4.1 ICT

The strategies of the government as set out in the PASDEP document[200] include: promoting human resource development in the ICT field; mainstreaming the use of ICT in all sectors of the economy, in the administration of government, and in the education system; developing the necessary telecommunications infrastructure; promoting research and development through ICT; and creating the enabling legal and regulatory framework. In addition to achieve having multiple access to international submarine cables and expanding improved telephone and Internet connectivity down to the each of the last administrative levels, the following targets are planned to be accomplished in the five year period[201]:

- A fully digitized national transmission microwave and switching system
- 4,000 km of fibre-optic cables along the major National Route 13 and 122+ interconnecting stations.
- 1,200+ broadband Very Small Aperture Terminals (VSAT) nationwide.
- A broadband multimedia Internet Protocol/Multi Protocol Label Switching (IP/MPLS) core network with footprints all over the nation; and a broadband Internet network with points of presence nationwide.
- Wireless Code Division Multiple Access (CDMA) to 5,000 rural villages.

The Ethiopian Telecommunication Agency (a regulatory body), the Ethiopian Telecommunications Corporation (the service provider), the Ethiopian ICT Development Agency, the College of Telecommunications and Information Technology, and the private sector work hand in hand in the implementation and realization of the strategy.

5.4.2 Telecommunications

[200] MoFED 2006, p.141
[201] Ibid.

The telecommunication infrastructure that countries build is necessary for any ICT enabled institution to conduct its day-to-day activities. The quality and geographic coverage of these infrastructures are essential to government institutions, international and local NGOs, business communities, higher institutions and research organizations and citizens to improve their productivity and services. The Ethiopian Telecommunications Agency (ETC) is state-owned and the sole telecommunication services provider. ICT has great potential that can be used in services such as health, education, distance learning, meetings, and workshops. Its importance is great especially in terms of cost and time savings compared to logistics and road infrastructure. A number of projects have been accomplished with the expanding infrastructure of the ETC. Currently, state-owned networks based on VSAT (Very Small Aperture Terminal) architecture are used. They include Woreda-net (district connectivity), Agri-net, School-net, and Health-net and are being used in their respective offices to communicate with each other and with the regional and federal government to facilitate the activities which curbs the above limitations of time, space, and cost. School-Net is a one-way link that enables schools to provide audio-visual lessons. Agri-Net supports agricultural research institutions in communicating and collaborating with each other. Woreda-Net is a net work of "Woredas" (districts at fourth administrative level) is used for building democracy and good governance. On top of building the telecommunication infrastructure, the objectives of ETC within the period 2004/05- 2009/10 include[202]:

- To reach 100 percent in the provision of telecommunication services in rural areas within 5 km.
- To increase the number of Internet users from 17,000 to 193,000 people
- To increase the teledensity of mobile subscribers by 8.1 (or 6.76 million people)
- To increase the teledensity of fixed line subscribers by 3.87 (or 3.23 million people)
- To be able to provide the ICTs to be used in the agriculture, social services, e-commerce systems, and industry sectors.

Telecommunications is one of the big infrastructures in Ethiopian. It was introduced in 1894 and is one of the oldest in Africa. It is a public sector body providing the sole services in fixed, mobile, Internet and data communications. The Ethiopian Telecommunications Corporation (ETC) is expanding its infrastructure and is expected to reach rural areas to

[202] MoFED 2006, p.140

achieve the Universal Access Service (UAS) and narrow the digital divide between the urban and rural households. However there is criticism, mostly from the ITU, that the type and quality of telecommunication services are poor, and the number of subscribers in the telecommunication service (fixed, mobile, and Internet) is very low, compared to other countries in sub-Saharan Africa. The ITU for example ranked Ethiopia at 147 out of 154 countries, with an ICT development index (IDI) ranking of 1.03 based on the 2007 estimates in its 2009 report with such component parts that constitute the IDI as access, use, and impact as shown in the table below (ITU 2009, pp. 13,22,91-95).

Indicators	Estimates 2007[203]
Access	
Fixed telephone lines per 100 inhabitants	1.1
Mobile cellular subscriptions per 100 inhabitants	1.5
International Internet bandwidth per Internet user(bit/s)	842
Proportion of households with computer	0.2
Proportion of households with Internet	0.1
Use	
Internet users per 100 inhabitants	0.4
Fixed broadband subscribers per 100 inhabitants	-
Mobile broadband subscriptions per 100 inhabitants	-
Impact	
Secondary gross enrolment ratio	30.5
Tertiary gross enrolment ratio	2.8
Adult literacy rate	47.5

Table 0-1 The ICT development indicator for Ethiopia

(Source: ITU, 2009, measuring the information society: the ICT development index)

However ETC has announced that it is going to increase the number of subscribers in fixed line, mobile, and Internet. It has also introduced reduced subscription tariffs[204] (see ETC website) to fixed and mobile telephones as well as Internet broadband and dial-up services.

[203] A hyphen indicates that the amount is nil or negligible
[204] ETC website: http://www.ethionet.et/services/fixedtariff.html, retrieved September 16, 2009.

5.4.3 Strategies of the CSA

In its medium term strategic plan, the CSA set out a number of objectives to be achieved at different times. Some of the objectives include:

- Increasing the number of surveys conducted.
- Making statistical reports at the "Woreda" (district) level.
- Reducing the delay in reporting by narrowing the time between data collection and dissemination gradually to regular and other surveys as planned.
- Enabling all branch offices to do data entry and data cleaning of some surveys.
- Creating data bank-to-main sectors in the CSA.
- Connecting all branch offices with a wide-area network (WAN).
- Improving the national statistical system by preparing standards such as concept definitions to be used in the "Woreda-Net" knowledge management initiative and all socio-economic sectors.

CSA can effectively implement its strategies if some of the strategies of the government in ICT and telecommunication are also implemented as specified in their document. For example, CSA may want to use mobile data collection in order to reduce the delay in reporting (one of the objectives above) and maintain some of the data quality dimensions in the IMF's DQAF. Enumerators may collect data using portable PCs (laptops, PDAs, UMPCs, mobile phones, etc) and send the collected data via a reliable telecommunication medium. This will become possible if CSA's internal ICT systems and the expertise of employees in this field (the IT and data processing department) are strengthened. When this is done in line with the materialization of the goals set out by other sectors such as the Ethiopian Telecommunication Corporation (ETC) and Ethiopian electric power corporation (EEPCO), it will have a beneficial effect. The ETC has a goal of expanding the telecommunication infrastructure both in type and quality of services, with greater coverage in the rural areas. The Ethiopian Electric Power Corporation's (EEPCO) aims to electrify most rural areas with hydropower, solar power, and wind mills.

One of the objectives of the ICT Department at CSA is to effectively implement the strategies set out by CSA. Some of the activities of the ICT department include the management of ICT

(hardware, software, and communication technologies), giving IT related training to staff in different departments and the selection and recommendation of appropriate ICTs that are used by different departments to support the statistical process to collect, capture, process, analyze, store, archive, and disseminate data and information. Troubleshooting of hardware and software problems both at the head office and branch office is done by the ICT department. The integration of all operational computers into the local area network (LAN) enables CSA to constantly update the antivirus software on all PCs to protect the software, hardware, networks, and data and information from risks, threats, and vulnerabilities. It also facilitates file and printer sharing, handling backups and usage of Internet over a dedicated broadband line. Part of the security measures include protecting the database and other statistical data and information by preventing unauthorized access, provision of usernames and passwords to all staff of the CSA who connect to the central server, and taking periodic onsite and offsite backups. In addition the department is in charge of the development of databases and websites usually with the involvement of external stakeholders for technical and financial assistance. The conversion of cartographic maps of primary sampling units, peasant associations, districts, zones, and regions are converted into digital form and stored in the GIS division of the ICT department. Upgrading its resources with the changes taking place both at national and global levels also improves the overall statistical system in order to meet the demands of stakeholder through the provision of quality data and information. This of course requires sound telecommunication and ICT infrastructure at the national (such as ETC) and institutional (such as CSA) levels that support the organization in collecting data, gathering this collected data from either the field or field branch offices, and disseminating the information to users over the World Wide Web using the Internet through its website and other electronic media.

5.5 ICT Infrastructure of CSA

The head office has two buildings that are connected by a fibre-optic cable. The computers in the two buildings are networked by a LAN using an unshielded twisted pair (UTP) with Star topology. The communication protocol used is TCP/IP. The LAN is served by a file server, bridges, hubs, and routers. Other servers include a web server, database server, mail server, and domain controller. There is a server dedicated to a firewall to protect the computer network of CSA from outside attacks such as viruses and unauthorized access. ETC is the

sole public enterprise Internet service provider (ISP) in Ethiopia and CSA is provided with the leased line Internet access broadband of 256 KB based on monthly cost of service fee. Professional and some technical staffs have access to the Internet using the LAN server which is protected by a firewall. In addition Norton antivirus corporate edition software is installed on the server and is periodically updated to protect the computers in the LAN.

There are a number of ICTs in use by CSA staff that supports the various statistical processes from questionnaire design to the collection, processing, storing, analyzing and disseminating of statistical data and information. The following table shows the lists by their primary use. PCs and some software, mainly PCs, and Microsoft application and systems software are used both at the head office and branch offices.

Category	Items	Use
Hardware		
	computers (desktops)	Generally used for data entry, data management, programming, analysis, report writing, documenting, etc.
	computers (laptops)	Analysis, presentations, meetings
	PDAs	Few are used in field data collection in price survey for selected areas (pilot).
	Servers	Used for storing records in databases, for web applications, exchange, DNS.
	Plotters	Used in graphics such as maps of EAs and other higher administrative levels.
	Digitizers	Digitize maps of EAs and other administrative levels.
	Photo Scribe PS900 Series optical mark scanner	Data capturing for the 2007 housing and population census.
	Scanners	Scan official documents.
	Photocopiers	Photocopy official documents.
	Printers	For office use and to print draft reports.
Software	Application software	
	Microsoft Word, Excel, Access, and PowerPoint	Prepare different survey documents (questionnaire designing, training

		manuals, reports, etc)
	Statistical	
	SPSS	Data manipulation and analysis.
	Harvard graphics	
	STATA	Data manipulation and analysis.
	Data management	
	Nesstar	Publishing survey data and information on the web for online analysis.
	IHSN	Archiving of datasets and survey documentation, and dissemination using CD-ROMs.
	Data processing	
	CSPro	Window based: Data capture, data cleaning, data manipulation, and tabulation.
	IMPS	Dos based: Data capture, data cleaning, data manipulation, and tabulation.
	System Software	
	Operating systems	Microsoft windows 2000 and XP are installed and used in different PCs
	Server Operating systems	Microsoft windows 2000 server platform used in different servers.
	Data base	
	MS-SQL server	o Ms-SQL server 2000 was used to build the producer and retail price RDBMS. o MS-SQL server 2008 enterprise edition is being used to build RDBMS for other surveys.
	EURO Tracer	Used to make a data base of trade and services.
	DevInfo	A database of socio-economic indicators that are helpful to stakeholders in tracking progress towards meeting the MDGs
	Personnel database	A database built using MS-Access and VISUAL BASIC for the human resource management department.
	Information system	
	Arc/GIS	Used to prepare and display information on a map by pairing surveys and censuses information and maps.
Communication	Presentation & media	
	Projectors	Used in workshops and trainings at the head office.

	Radios and TVs	CSA uses State owned radio and television to disseminate survey information. It is also used to raise public awareness to get cooperation from respondents when surveys and censuses are conducted.
	Networking	
	Hubs	Connect PCs in the LAN to improve data transmission.
	Routers	To share Internet for networked computers, to connect networks of computers.
	Network cables (by type)	Unshielded twisted pair (UTP) category 4 is used in LAN; Fibre-optics is used to connect the LANs in the two buildings.

Table 0-2 Major ICTs and their usage in CSA

(Source: compiled by the author)

During pre-data collection most of the activities performed include consultation, questionnaire design, printing of questionnaires, training of trainers, dispatching of trainers and questionnaires to branch offices, and training of enumerators. In addition to system software, application software such as Microsoft Word, Excel, and printers are used to perform the above activities. Exchange of ideas and draft questionnaires between various internal stakeholders takes place through the Local Area Network (LAN).

During the data collection stage the mode of data collection to most of the socio-economic surveys used is a face-to-face PAPI mode. Data collectors interview respondents and fill in the forms using pencil with the aid of guiding manuals. Questionnaires are designed first by SMS in collaboration with the DP, methodology department, and field operation staff. The DP department develops and tests the data entry system based on the final draft of the questionnaire and reports back any errors or inconvenience encountered. In the PAPI mode no ICT is used. CAPI mode of data collection using a total of 138 PDAs (hp iPAQ with backups) is used to collect consumer price index (CPI) and producer price index (PPI) data for selected market places in five branch offices. The application is developed using Microsoft Excel with some controls on minimums and maximums. The local Amharic (the official working language of Ethiopia) font is used in the application. Before the

implementation of the CAPI mode, a comparison was made with the PAPI mode with respect to data entry errors (using three PDAs by enumerators and three keyboard entry by data entry operators) to decide whether PDAs can be used for data collection. The comparison showed a tolerable error rate, giving the green light to the use of PDAs. It is now possible to receive data collected using PDAs within two days after collection is completed. This gives rise to some improvement in the *accuracy and reliability* (reduced error) and *serviceability* (timeliness*)* data quality dimensions of the IMF's DQAF. However, problems such as battery life (only two hours), weak computer literacy and lack some mathematical skills of enumerators, and problems related to finding the right type PDA (solar powered) were reported. For other surveys, mobile data collection using PDAs or other portable PCs, assessment needs to be done regarding the overall benefit of mobile data collection taking into consideration the various data quality dimensions as mentioned in the previous chapter and cost savings in the long run. The process of collecting survey data is more or less similar to most surveys but each survey has different degrees of complexity. The CPI and PPI surveys are far simpler in terms of complexity of the questionnaire compared to the DHS and HICES questionnaire for example.

Web-based data collection is not practiced by the CSA. Computer-assisted self interview (CASI) mode can be applied, given current Internet penetration, to industry and surveys. Questionnaire forms that are attached to the database can be uploaded for self-completion by respondents in the respective industries or establishments, or sent by e-mail. The completed forms are then sent back for further processing. This is possible if there is a reliable underlying national telecommunication infrastructure and the CSA's own ICTs infrastructure. This improves the *timeliness* data quality dimension by reducing the time it takes to fetch data from each branch offices by car in addition to the cost of logistics. This is in line with the basic goal of CSA in providing reliable and timely data and information.

However some experiences show that there are some barriers to using PDAs[205]. A common problem would be in the rural areas where there is no electricity. The majority of data collectors live within the delineated enumeration area, which is far from power sources, which may hinder their use unless they have extra batteries which last longer. Recent PDAs

[205] MoH 2007

come with battery that can last long enough to conduct interviews for a number of households depending on the quantity and complexity of the type of survey. Another problem with PDAs would be the font size, which makes them difficult to be read. The PDA should also support the local language alphabets, as most survey questionnaires are designed in local languages. Portability and secure transmission lines could be other barriers that may slow down activities. Maintenance and repair of laptops or PDAs in case of failure is also another issue, as it is either expensive or the service might not be available in the branch office. However problems related to power and communication mediums such as the Internet could be overcome if what has been planned in the second phase of the PRSP of Ethiopia (i.e. PASDEP) is effectively implemented in the specified time. Regarding rural electrification, the Ethiopian electric and power corporation has an ambitious plan under its Universal Electrification Access Program (UAEP) to cover 50 percent of the entire population by 2009/10, whereas the Ethiopian Telecommunication Corporation plans to expand both the type of services and number of users to increase access to telecommunication within 5 km to 100 percent, increasing the teledensity of both fixed and mobile telephones, and increasing Internet subscribers by the year 2009/10[206]. According to EEPCo[207] the number of electrified towns increased from 16 percent in 2004/05 to 1658 (22 percent) in 2007. ETC has a plan to increase the telephone coverage to "90 percent of the nation by January 2010" and in 2009 it has the following number of subscribers as stated by the Capital newspaper[208] "*…mobile subscribers: 4,051,703; number of fixed-line subscribers: 902,955; CDMA subscribers: 56,310; Internet and data subscribers: 74,557*." The use of mobile data collection and transmission therefore, to some extent, goes hand-in-hand with the development of infrastructures in the telecommunication and power sectors in the country.

The Post-data collection process in the CSA includes registering questionnaires received from field, documentation, data editing and coding, data entry, computer data cleaning, data preparation, and tabulation, data basing, data archiving, and dissemination. In addition to the Windows operating system (2000, XP, Server), various other communication tools, hardware and software are used in the post-data collection stage for data entry, data management, programming, analysis, report writing, documenting, presentation, digitizing primary

[206]MoFED 2006, pp.l39-140
[207]EEPCo website: http://www.eepco.gov.et/brief.html, retrieved 15 September 2009.
[208]Kirubel Tadesse 2009

sampling unit (or enumeration area) maps, printing, archiving and uploading survey information. IMPS is developed by the International Program Centre of the US Census Bureau. CSPro on the other hand has been developed by the International Program Centre of the US Census Bureau in collaboration with Serpro, and Macro International. These data tools are used in the development of data entry systems, a system to clean or edit the captured data, and tabulation[209], which is used in the report. In addition to the above software, SPSS, Excel, and Word are also used. SPSS is mostly used in data manipulation and analysis whereas Excel sheets and Word are used to prepare the survey report.

The ICT department, in addition to taking care of the IT management in CSA, is responsible in the storage, retrieval, and dissemination of information. It provides raw data to stakeholders upon filling in a request form. There are database servers, Web servers, mail servers, and domain name system (DNS) servers. The staffs in the ICT department install and upgrade new operating systems and software, troubleshoot desktop PC, server, or network problems, both at the head office and branch offices, recommend appropriate ICTs and participate when software is developed by developers outside the CSA, and give IT-related training to CSA staff. The websites and databases are also maintained by the ICT staff. Survey documents are also converted to PDF using the Adobe Acrobat PDF maker software on completion of all surveys. In addition, metadata is created for each completed survey and the dataset, along with its survey documents, are archived using Integrated Household Survey Network (IHSN) micro data management toolkit software. These are then uploaded to the website for dissemination. Survey documents are also published on CD-ROMs. Stakeholders can access the website for specific digital survey information from the data catalogue. This is a national data archive which contains a catalogue of surveys conducted by CSA over a number of years.

This catalogue provides survey information such as a general overview of the surveys, metadata, and various other documents (technical information, reports, and training manuals). This improves the *Accessibility* data quality dimension of the DQAF of the IMF, mainly by improving the *data accessibility* and m*etadata accessibility* elements. The catalogue also has a search facility, which lists all surveys containing key words. The website contains static

[209] A detailed description of these activities with flow charts is described in the next chapter

files such as the yearly statistical abstract, news bulletins of the CSA, and organizational information in the form of PDF files. There are two online databases on the website of CSA. It is also possible to do online queries using both databases: the EthioInfo and Price databases. EthioInfo is adapted from DevInfo developed by UNICEF and is used in many UN member countries. The main purpose of EthioInfo is "*to organize, and present data in a result based environment with unique features linking to strategic monitoring and evaluation of policies such as MDG, National Poverty Reduction Strategies...EthioInfo facilitates data sharing and indicator harmonization at global, regional and country level by making statistics available to a wide audience.*"[210] It is of great use to stakeholders such as planners, line ministries, research institutions, higher education institutions, the media and UN affiliated agencies. The price database is developed using SQL 2000 and contains consumer and producer prices. The database has a search facility which displays prices of food and non-food items for any year, month, and market places.

The ICT department has a GIS section. This section creates digital maps using digitizing software at different Government administrative levels as well as the enumeration area (EA) that is used as the lowest sampling unit which CSA uses. It also uses ArcGIS and ArcView software to integrate the different socio-economic and census data and information to disseminate using a map. Stair and Reynolds[211] define a geographic information system (GIS) as "*a computer system capable of assembling, storing, manipulating, and displaying geographically referenced information, that is, data identified according to its location.*" The GIS section in the ICT department of CSA so far has disseminated the 2007 population and housing census and the 2000 agricultural enumeration survey, linking geographically disaggregated tabular statistical outputs on their respective maps. Disseminating information using GIS technology enables decision- and policy-makers to visualize data and information and spatially compare the levels of poverty at different geographic levels. It also allows visual monitoring of the outcomes of anti-poverty programs by looking at the different indicators such as income, access to basic services (such as safe drinking water, health, cooking fuel, market, telephone, radios, gross and net enrolment ratios in primary, secondary levels, etc), and total population size. Representing poverty-related data and information on a map is therefore important to plan, monitor, and execute the PRSPs and various anti-poverty

[210] CSA: http://www.csa.gov.et/di5web/content/information/product/product%20en.htm
[211] Stair & Reynolds 2006, p.480

programs which ultimately enable sustainable socio-economic development both in the rural and urban areas. This requires improving the quality of data and information generated by CSA. GIS technology improves the *serviceability* and *accessibility* data quality dimensions in the IMF DQAF.

5.6 Conclusion

In this chapter the roles of CSA and other external stakeholders was discussed in relation to the PRSPs and data quality. Data quality is also discussed in relation to the various ICTs used and the different activities performed by CSA during the pre-data collection, data collection, and post-data collection stages to collect, process, store, archive, analyze, and disseminate data and information. The next chapter provides a detailed description of how socio-economic data is generated as well as the various resources used in most of the household surveys.

CHAPTER SIX

ANALYSIS

6.1 Introduction

This chapter gives a detailed description of how quantitative survey data is generated by CSA. It describes the flow of activities (flow charts) that are used to produce data and information, followed by a description of this flow that discusses the inputs, processes, players and outputs involved at each stage of the statistical process. The chapter ends with the conclusions drawn.

The section on analysis in this chapter is based on what was observed in CSA. Flow charts and descriptions are used to interpret how data and information are generated as well as the different interactions between employees. The flowcharts and the descriptions that follow, which are based on observations and direct involvement in the real job both at the head office and branch offices, are similar to most of the household surveys conducted by CSA including those used in poverty analysis. The observations include different activities, interactions, and use of ICTs and other resources during the pre-data collection, data collection, and post-data collection stages of the statistical process.

6.2 Flow Charts

In this section the flow of various activities or processes that are used to generate socio-economic data and information and the various stakeholders involved in the integrated household survey project (for example WMS, HICES, agricultural, and most other household surveys) are shown using flow charts in four parts. The first figure (Figure 6-1) shows the general flow chart of activities at all levels, both at the head office and branch offices, used in most of the surveys. The second figure (Figure 6-2) shows the general flow chart of activities in the branch office, containing details of Figure 6-1. The third figure (Figure 6-3) shows the

general flow chart of activities in the data-processing department at Head Office. It describes the various post-collection processes that take place after data and questionnaires have been received from the field. The fourth figure (Figure 6-4) and fifth figure (Figure 6-5) show the detailed flow chart during data entry and data cleaning in the data processing department. They also show data and questionnaire flows and the interactions between the different departments within the CSA during the post-data collection stage.

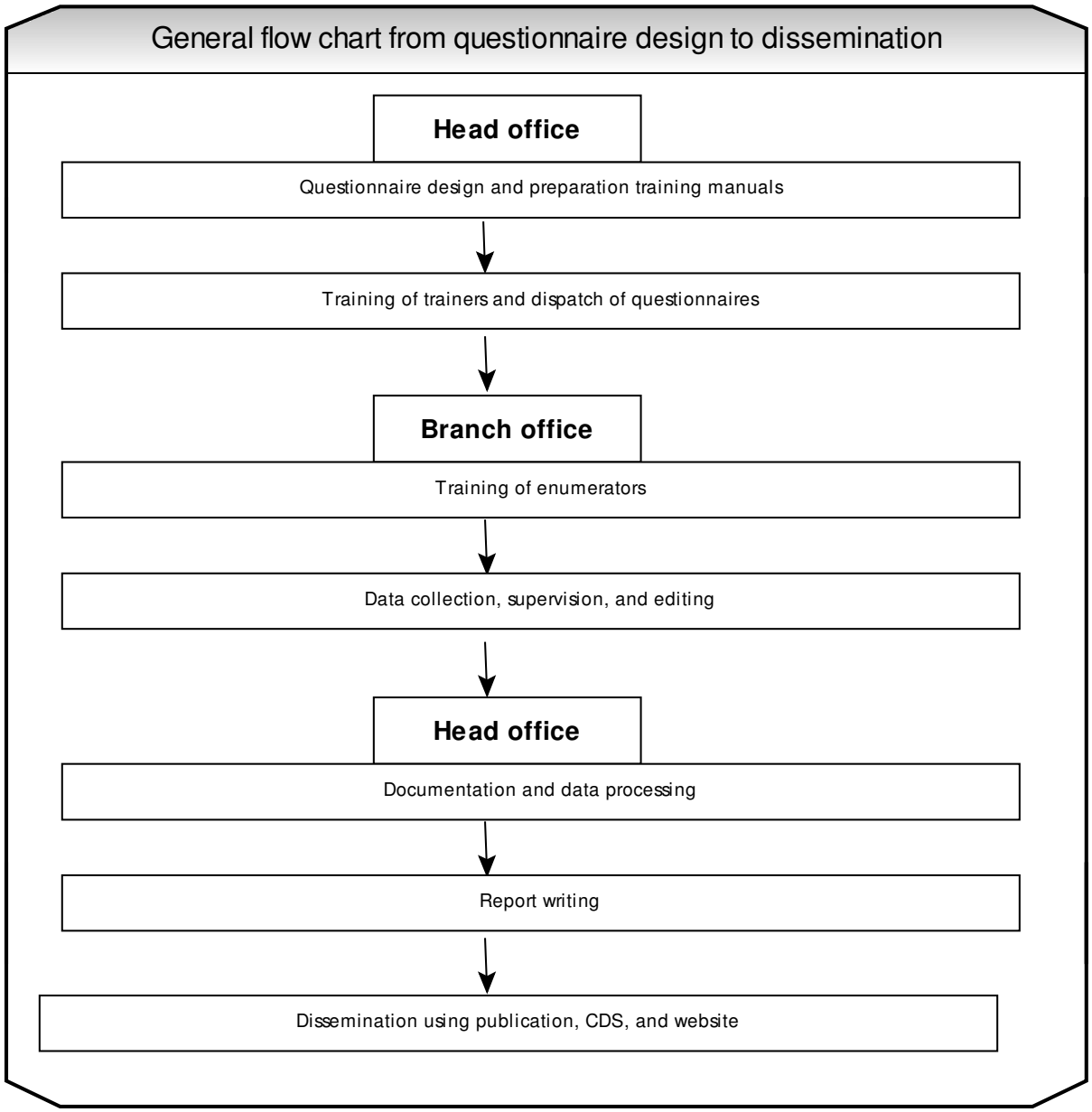


Figure 6-1 General activities at Head Office and the branch offices.

(Source: Author's illustration)

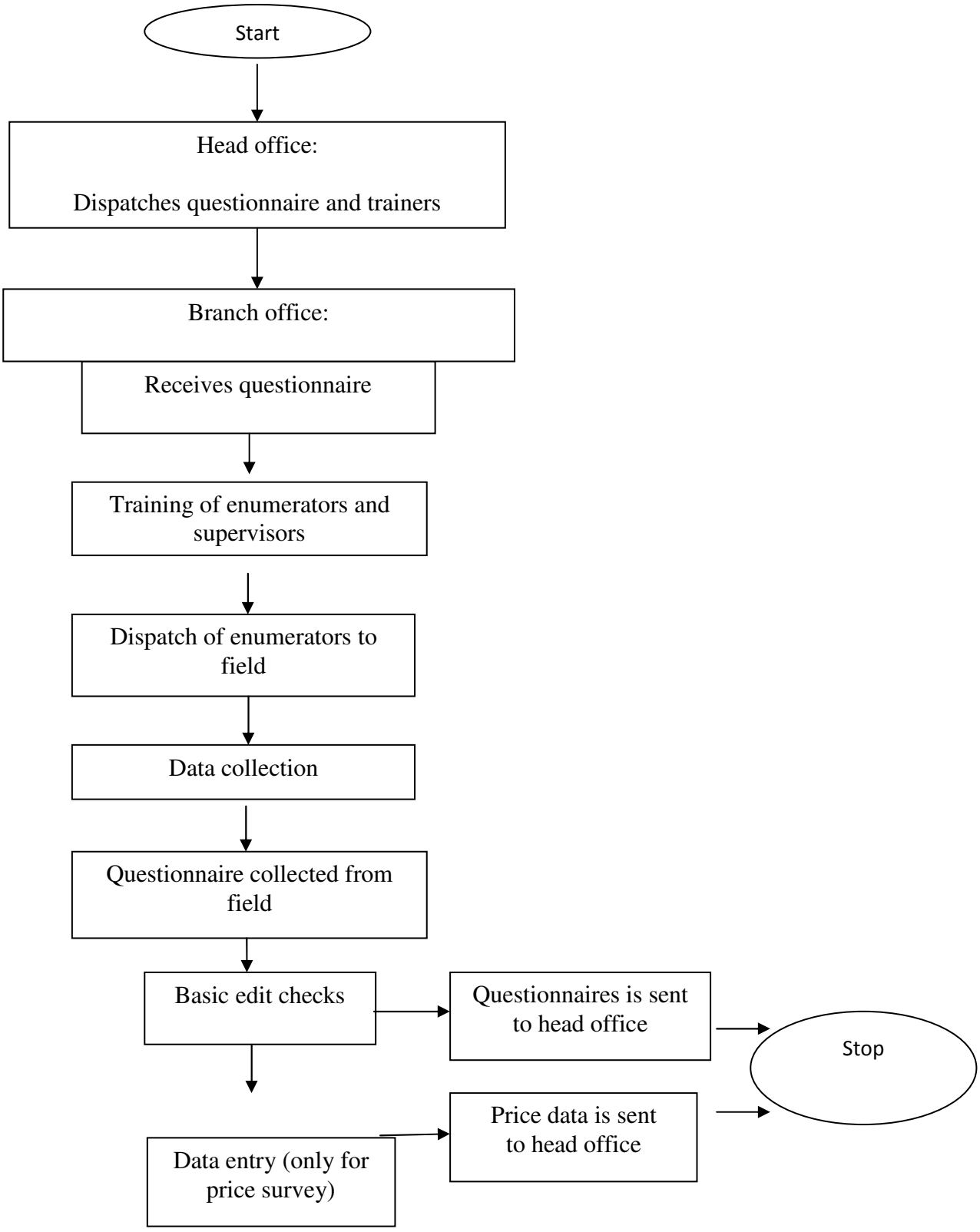


Figure 6-2 Detailed questionnaire and data flow in the branch offices

(Source Author's illustration)

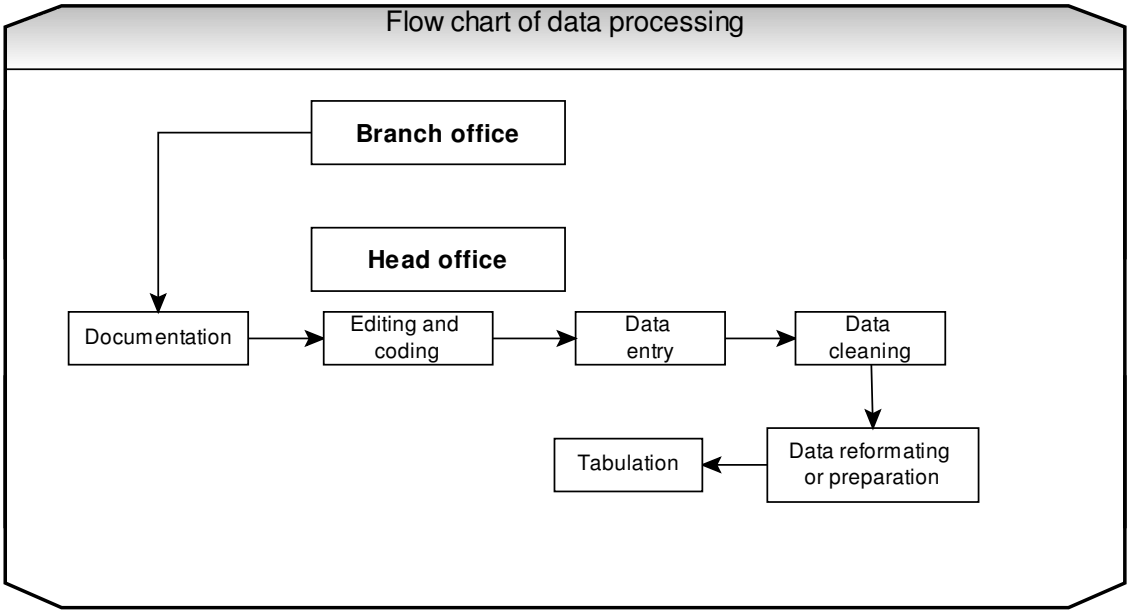


Figure 6-3 General questionnaire and data flow in the data processing department at the head office.

(Source: Author's illustration)

Except for the price survey, all completed questionnaires of other surveys are sent to Head Office. Here they are received and registered by the Documentation Office of the Data-Processing Department after being notified by the Field Operations Department. From documentation they are sent to the editing and coding office, then received by the data entry managers to be distributed to data entry operators for data capturing; the soft copy then moves to the data cleaning office and data is cleaned, with frequent reference to the questionnaires. The cleaned data is then manipulated and reformatted to be used for tabulation. Data entry for price surveys is done at the branch offices, copied onto diskettes and sent to Head Office. First the documentation office is notified and then the diskettes are given to the programmer or system analyst in the systems and programming division in the data processing department for further processing, mainly cleaning and tabulation.

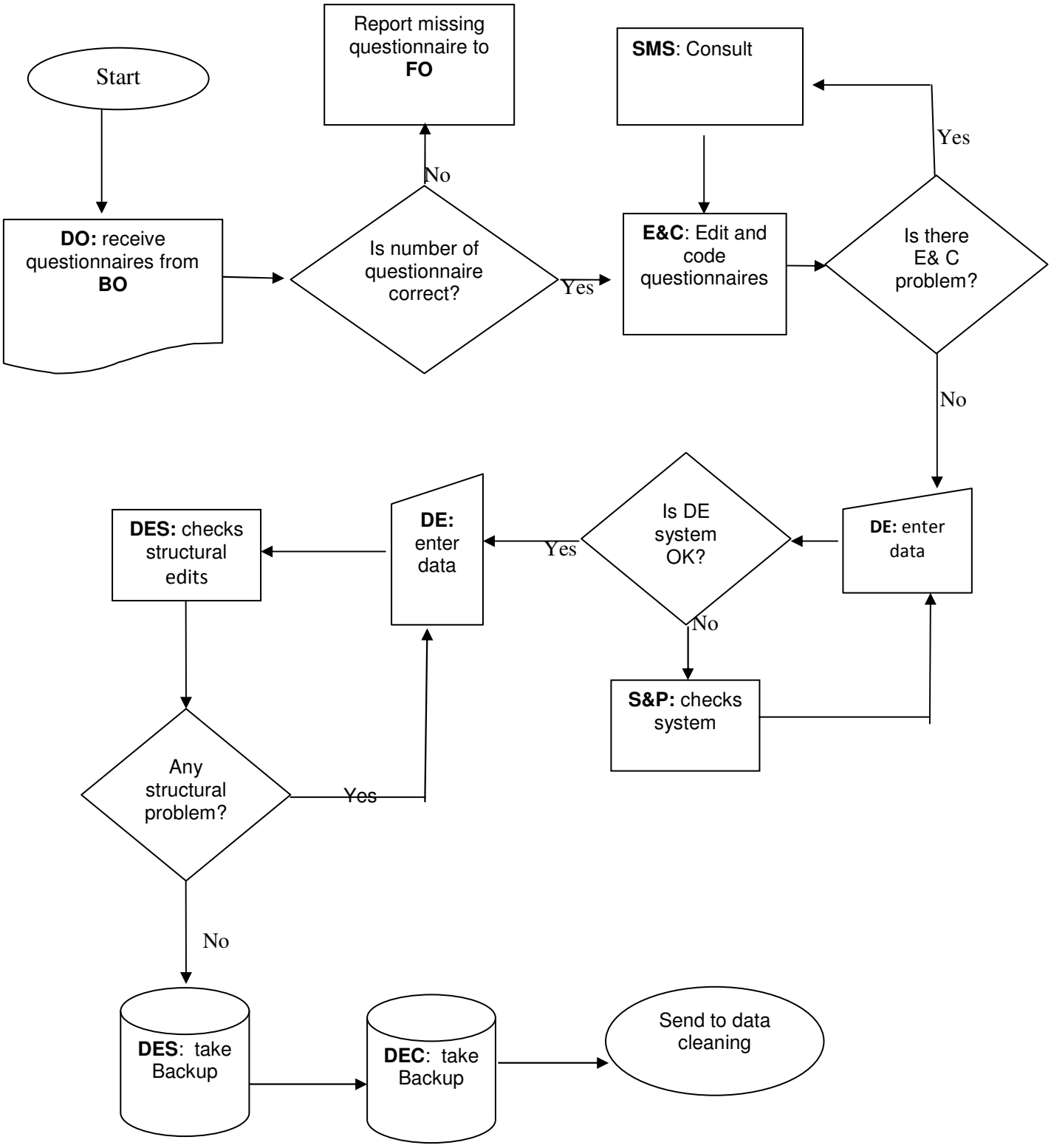


Figure 6-4 Flow chart of activities during data entry

(Source: Author's illustration)

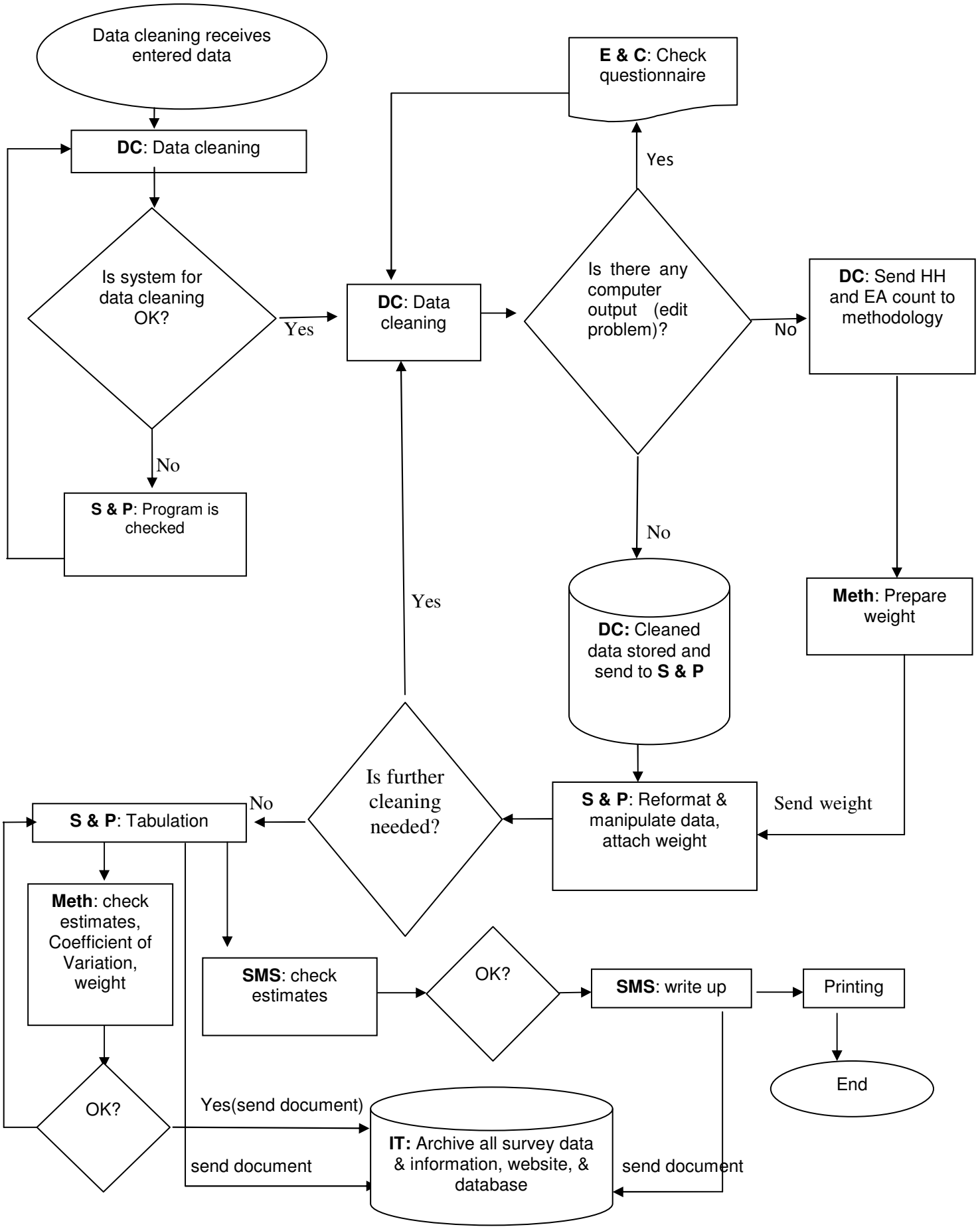


Figure 6-5 Flow chart of activities during data cleaning

(Source: Author's illustration)

Note to Figure 6-4 and Figure 6-5:

DC: data cleaning, *DE*: data entry, *DEC*: data entry coordinator, *DES*: data entry supervisor, *DP*: data processing, *E & C*: editing and coding, *IT*: Information technology department, *Meth*: methodology department, *SMS*: subject matter specialists (experts in the line departments such as Natural Resource and Agricultural Statistics, Industry, Trade and Service Statistics, Household Budget Welfare Manufacturing and Prices, etc), *S & P*: systems and programming division.

6.3 Description of Activities in Generating Data and Information

Below is a description of the different inputs, participants (players), process, and output both at head office and branch offices that produces the data and information as described in the above flow chart. The *heading* refers to the location where the activities are performed, whether at head office or branch offices, and the divisions within the department. *Inputs* refer to the human and material resources required to perform the activities in the indicated department/division. *Participants* are those people in the department/division who are responsible for making and executing decisions, and the activities at each level. The *process* is the actual activities performed and *outputs* are what are expected from participants in their respective departments/division that are used as inputs to the next phase in the flow. Different ICTs are used in some of the steps in the system. Possible *ICTs* that may improve data quality are suggested at each stage in the flow. The outputs from each step of the statistical process are obtained from the inputs at the previous steps, the human and material resources and the actions and interactions between players.

The description is arranged in three parts, namely, *pre-data collection, during data collection,* and *post-data collection*. The pre-data collection phase includes the first five steps and they include stakeholders meetings, questionnaire design, training of trainers, dispatch of questionnaires and trainers and enumerator training in the branch offices. During data collection, only the data collection step is included. The post-data collection step includes the remaining fifteen steps, of which the first two are taken at the branch office, the next twelve at the head office and the last taken by MoFED. These steps are: branch office editing and coding, branch office and head office editing and coding, data entry operation staff, systems

and programming, the data entry process, systems and programming, the data cleaning, methodology service, subject matter specialists (SMS), systems and programming, methodology department, subject matter specialists, printing, ICT department, and MoFED.

PRE-DATA COLLECTION

Head office - meeting of stakeholders:

Input: Past surveys reports summary, new questionnaire design formats for discussion

Participants: Different CSA staff and stakeholders, depending on the nature of the survey includes from different line ministries, non-governmental and international organizations (for example CSA, MoFED, Ministry of Labour and Social Affairs, Ministry of Rural and Agricultural Development, NGOs, WB, etc…). CSA SMS staff and other stakeholders interested in the survey.

Process: Stakeholders evaluate and discuss lessons learned from similar past surveys conducted. Discuss new and additional indicators and features for future survey about their budget, formats, etc…

Output: budget and technical assistance, indicators, formats of questionnaires.

Head office - questionnaire design:

Input: Indicators and formats of questionnaires, PCs, laptops, MS Word and Excel, Windows 2000/XP

Participants: SMS, DP, and Methodology staff of the CSA.

Process: SMS design the questionnaire and then discuss it with DP staff who will be using the questionnaire to develop data entry and processing systems. At the same time they also prepare table formats that will be used in the report, training manuals for enumerators and supervisors and a manual that will be used for editing and coding. DP staff are involved only in the design of the questionnaire and report problems encountered, related to ranges, skips, consistencies, presence of identification particulars (IDs) for each questionnaire, presence of linking and basic variables in each form within the questionnaire, which is important in the cleaning and tabulation format, and checking overall compatibility of the questionnaire design with the software that is going to be used in data entry, cleaning, and tabulation. These feedbacks are incorporated and then the final version of the questionnaire is printed. The methodology staffs verify the questionnaires whether columns that are used in the sample selection, such as the geographic IDs and household identifications, are included and that the codes used follow the standard.

Output: Questionnaire, training manual, codes, edit specification for manual and computer editing, tabulation formats

*ICTs***:** PCs, laptops, Operating system, MS Word and Excel, CSPro, IMPS are used in the design of questionnaires and tables, and data entry systems.

Head office - training of trainers:

Input: Questionnaire, training manual, codes, trainees.

Participants: SMS in the line department who participate in the design of the questionnaire and manuals give the training. Other staffs from the line department, other departments, and branch offices (both professionals and sub-professionals) who will be attending the training are selected.

Process: Face-to-face training is given on the survey to professionals and sub-professionals at the head quarter by line department SMS who participated in questionnaire design and manuals. After the training, mock interview is conducted in the field to test the questionnaires, manuals, and codes. Discussion is held after the mock to get feedback and further improve the questionnaire, manuals and codes.

Output: Final questionnaire, training manuals, and trainers.

*ICTs***:** Projector, PCs and speakers improve the audibility and visibility of the training, questions and discussion to all participants. This is especially important where training is conducted in big halls for a large number of trainees. Depending on the number of trainees and the complexity of the survey, video conferencing and VOIP can be used. For example in the 2007 Population and Housing Census data entry was done using scanning technology. Therefore careful handling and correctly filling the questionnaires was required to reduce rejection of forms during scanning. Video conferencing was used from the head office to trainers of data collectors on how to handle questionnaires and fill in the forms carefully. This can also be applied to other surveys such as the WMS and HICES.

NOTE: Due to lack of halls that accommodates a large number of trainees in the head quarter, CSA usually rent halls. Some times more than one hall is rented depending on the number of trainees.

Head office - dispatch of questionnaires and trainers:

Input: Final questionnaire, training manual, codes, trainers, vehicles,

Participants: Field operation staff, SMS, trainers, finance office, drivers.

Process: Trainers are assigned to branch offices to give training, questionnaires and manuals are packed, finance office gives daily allowances to trainers for field work and finally trainers and questionnaires drive to the assigned branch office. In each branch office enumerators are also expected to arrive for the training. Recruitment of enumerators is also done when necessary.

Output: trainers arrive at the Branch office with questionnaires and manuals.

*ICTs***:** PCs, laptops, MS Word and Excel, Windows 2000/XP are used in record keeping of vehicles, number of questionnaires, trainers, enumerators, and supervisors who are going to be involved in the survey.

Head office - Branch Office enumerator training:

Input: trainers, questionnaires, manuals and other measuring instruments such as balances, meters, vehicles.

Participants: trainers, head of branch office and technical staff, enumerators, branch office finance office, drivers.

Process: Enumerators receive their allowances. Manuals and questionnaires are distributed to enumerators. Face-to-face training is given to enumerators in more or less the same way as it is given at the head quarter. Sharing of experience is held while enumerators raise questions they faced in the past and how they dealt with it. Feedbacks are discussed after mock interview. SMS in the head office are always contacted to resolve technical problems encountered during the training. The SMS then inform any correction whatsoever to other branch offices to maintain uniformity. Discussion with supervisors is also held separately on topics related to supervision (if it is not held at the headquarters). Finally drivers take enumerators to their enumeration areas.

Output: Final questionnaire, training manuals, measuring instruments, trained enumerator.

*ICTs***:** So far ICTs are not used. Projector, PCs, and speakers improve the audacity and visibility of the training, questions, and discussion to all participants. This is especially important where training is conducted in big halls for a large number of trainees. Buying projectors to all the 25 branch offices may not be realistic specially when there are other priorities, but they facilitate the discussion by providing trainees visually appealing and intuitive understanding of *concepts, definitions, and techniques*. This improves the *methodological soundness* and consequently other related data quality dimensions. Alternatively, connecting the branch offices with Wide Area Network (WAN) would also be another option to connect directly with enumerators while the raining is given. Virtual private network (VPN) over the telecommunication infrastructure also enables to use voice over Internet protocol (VOIP) and Video conferencing using Internet. With these technologies sound and visual images from the source (e.g. head office) are converted to IP packets and converted back to sound and visual image in the destination (e.g. branch office) to discuss about difficult concepts and share the scarce human and material resources. Currently experts at the head office who design the questionnaires and prepared the manuals are always available to answer technical problems asked from each branch office using telephones. These explanations then are discussed by the trainers in the branch office. Providing first-hand information from head office to trainees to all branch offices supported by further explanation by trainers is more advantageous. Therefore using video conferencing during or at the end of the training session to clarify the most confusing technical problems, enables the statistical process to be standardised and improves the quality of data.

NOTE: This training is one of the most important as it is given to data collectors who are directly involved in interviewee-respondent paper form-based data collection. Enumerators usually come with first-hand experience regarding problems they encountered while collecting data at other surveys in the past. They usually report these problems and share their solutions. The training also acts like lessons learned in the past and provides valuable experience, both to enumerators and SMS, which builds the institutional social capital. Due to the lack of a hall in some branch offices, halls are usually rented. Sometimes more than one hall is rented depending on the number of trainees.

DURING DATA COLLECTION

Head office - branch office data collection:

Input: Final questionnaire, training manuals, measuring instruments, random tables, and trained enumerator.

Participants: From headquarters – SMS and field operation staff. From branch office – head, statisticians, technical staff, supervisors, enumerators, and drivers.

Process: Enumerators first delineate the enumeration area, list all households within the EA on the paper form, select sample households from the list, give appointments to selected households for interview, and collect data from the sampled households using paper form. One Supervisor checks five enumerators on average and checks the listings, gives random start number that helps to select required sample households for interview, assist enumerators when faced with problems related to filling the questionnaire, solve problems when respondent resist to give interview, deal with other problems related to security and resolve or reports to the head of branch office if it requires further consultation, fetch salaries from branch office for the field enumerators, provide other missing or additional input necessary when needed by enumerators (questionnaires, better measuring instruments, pencils, rubbers, etc). Statisticians perform spot checking in selected enumeration areas. SMS in the head office keep on-line to resolve unfolding technical problems encountered in the field. Field operation staff at headquarters gives logistic support such as dealing with any problem related to vehicles, dispatch more questionnaires and other inputs necessary for data collection.

Output: filled-in questionnaire

*ICTs***:** So far about 138 PDAs were used to collect the monthly consumer price index (CPI) and producer price index (PPI) to some selected markets. MS Excel is used to develop the data entry system using PC and laptops and dumped on the PDAs. Mobile data collection using CAPI instrument installed on portable PCs (such as PDAs or UMPCs) can be used for selected surveys. Data collection and data entry are combined as a single activity. Sometimes computer editing or data cleaning is also included depending on the type of survey. Some of the advantages of using the CAPI method of data collection include automation of entry process like sequencing, cross-checking variables and maintaining their ranges, automatic skips to columns or pages, and consistency check. If portable PCs have GPS incorporated, then it is possible to locate households who are selected for interview; this is especially important for new enumerators who are assigned to new enumeration areas in order to complete a survey started by another enumerator.

Portable PCs require literacy of enumerators to perform operations such as basic operating system, application and custom software, and basic troubleshooting. Therefore proper training to enumerators is important in order to gain the benefits of using mobile data collection. Moreover if the task of editing and coding is done during data collection, as in the case of some developed countries, then enumerators need to have proper editing and coding training. Data collectors may also face hardware problems in addition to its systems capacity. Battery life is also another problem due to no electricity in the field for charging the battery. The mobile data collection device could also be lost which means total loss of household data collected in the enumeration area if backup mechanisms are not in place or data is not sent

every day to head office over a telecommunication medium. Another problem from an organizational view is that if data collection for all surveys is done using portable PCs, it may affect some of the roles of data entry and other personnel in the data processing department and may even call for a change in organizational structures.

POST- DATA COLLECTION

Branch office - branch office editing and coding:

Input: filled-in questionnaire and vehicles, PCs, MS Word and Excel, Windows 2000/XP, CSPro, IMPS

Participants: Statisticians, technical staff, supervisors, drivers, and field operation staff at the head quarter.

Process: Statisticians, technical staff, and supervisors perform preliminary data editing and coding. Drivers collect questionnaires from enumerators in the field. Field operation staff at the head quarter gives logistic support like providing vehicles and drivers.

Output: Edited and coded filled-in questionnaire (preliminary)

*ICTs***:** PCs, MS Word and Excel, Windows 2000/XP, CSPro, IMPS are used. If portable PCs (such as PDAs, or UMPCs) are used during data collection, then the data from different sites can be dumped onto PCs at the branch office for verification and preliminary editing and send the system to the head office for further editing and cleaning. If all the detailed computer editing and cleaning is done at the branch office, then the system that is used at the head quarter has to be applied. This means that the part of the data processing system that performs the computer editing must be distributed to all branch offices and they can clean the data. This has also calls for changes in the organizational structure.

Branch office and head office:

Input: edited and coded filled-in questionnaire (preliminary) and vehicles, PCs, MS Word and Excel, Windows 2000/XP

Participants: supervisors, drivers, field operation staff at the head office, and documentation staff in the DP department.

Process: Supervisors and drivers report to field operation at their arrival. Documentation staffs in the DP department receive and document the questionnaires and store them in the documentation hall.

Output: Documenting edited and coded filled-in questionnaire (preliminary)

*ICTs***:** PCs, MS Word and Excel, Windows 2000/XP are used. If CAPI is used for data collection, then only the soft copy is registered for check-in as in the price survey, where data entry is done at the branch office, the data is sent to the head office using a storage media and the documentation office is notified. If PAPI is used, then automating the documentation of questionnaires using a database that serves both the

documentation unit and field operations department could be used to follow up activities and report missing files.

Head office – data processing, editing and coding:

Input: Documenting edited and coded filled-in questionnaire (preliminary), PCs, MS Word and Excel, Windows 2000/XP

Participants: Documentation staff, SMS, editors, head of editing and coding staff.

Process: At head office the filled-in questionnaires from the documentation office are distributed to editors for further manual editing and coding. Here fields that are reserved to be filled in at head office are coded, fields that were skipped but are important are compared and coded, new responses to fields are compiled and accordingly new codes are given to them, and other consistency checks are made. Finally the edited and coded questionnaire is documented.

Output: Edited and coded questionnaire

*ICTs***:** PCs, MS Word and Excel, Windows 2000/XP are used

Head office – data processing, data entry operation staff:

Input: Edited and coded questionnaire, PCs, MS Word and Excel, Windows 2000/XP

Participants: Documentation staff, SMS, data entry operations staff, head of editing and coding.

Process: Documentation staff register all the edited and coded questionnaires and make sure that they reach data entry operations staff. Data entry operation staffs receive the questionnaires, compare their number with the selected number of sample EAs (using geographic code book), assign a number of questionnaires to each data entry operator, make sure that the PCs are working, and check that entry programs are installed.

Output: Edited and coded questionnaire as in the geographic code book

NOTE: Data entry programs are prepared by programmers and system analysts of the systems and programming division in the data processing department.

*ICTs***:** PCs, MS Word and Excel, Windows 2000/XP are used

Head office – data processing, systems and programming:

Input: questionnaire or Edited and coded questionnaire as in the geographic code book, PCs, MS Word and Excel, DOS, Windows 2000/XP, CSPro, IMPS, C++, BASIC, LAN

Participants: Programmers, systems analysts and SMS

Process: Programmers or system analysts design the data entry system early, before data collection begins, using the questionnaire and test it again with filled-in or edited and coded questionnaires. PC based data entry systems that are capable of capturing the edited and coded questionnaires must be installed. In addition to capturing data

from the edited and coded questionnaires, the system performs range checks and skips, and enables data entry operators to access electronic data entered by other operators through a LAN network and perform verification (entering again). It also allows supervisors in the entry room to access the verified data through a LAN network to make structural edits which checks for completeness of the sampled households in each EA, lets them modify the data for correction, and make back-ups when done with all the corrections. The final system is sent to data entry operations staff and installed onto the data entry operators' PCs. In some surveys when data entry is well underway, the system is sometimes modified because of unforeseen situations in some of the variable's properties such as length, decimal places, inclusion of new variables, etc.

Output: Data entry system.

ICTs: PCs, MS Word and Excel, DOS, Windows 2000/XP, CSPro, IMPS, C++, BASIC, and LAN are used.

Head office – data processing, data entry process:

Input: Edited and coded questionnaire, code books, PCs, MS Word and Excel, Windows 2000/XP, Data entry system done using CSPro or IMPS, C++, BASIC, LAN

Participants: Data entry operators, data entry supervisors, data entry operation staff, programmers or system analysts and SMS.

Process: Each questionnaire is entered by two different data entry operators. The second entry is to verify the data entered by another operator and so improve the quality of statistical data. Each operator's PC is connected to a LAN so that exchange of data takes place between operators and with supervisors. The verified data is accessed by the supervisor and is checked for completeness, which enables the supervisor to find missing households and take corrective measures. Finally the supervisor backs up the verified and structurally edited data, which is then collected by data entry operations staff for further scrutiny in the next phase. Programmers and analysts stay close to data entry operators to modify the system if there is a problem. The SMS gives support when there is a subject-related problem. The overall activities of the data entry process are closely monitored by the data entry staff. Problems related to hardware, entry system, power, and supply of edited and coded questionnaires to entry operators who are already done.

Output: Data (verified and structurally edited)

ICTs: PCs, MS Word and Excel, DOS, Windows 2000/XP, CSPro, IMPS, C++, BASIC, and LAN are used.

Head office – data processing, systems and programming:

Input: Edited and coded questionnaire, data (verified and structurally edited), PCs, MS Word and Excel, DOS, Windows 2000/XP, a data cleaning system done with CSPro or IMPS, C++, BASIC, LAN

Participants: Programmers, systems analysts, data cleaning staff, and SMS

Process: The systems and programming division design data cleaning programmes that check for structural checks, skips and ranges, duplicates and inconsistencies, the count of households and EAs and other errors based on the edit specification provided by the SMS. The program enables to correct errors both on the questionnaire and the verified data. The program is checked using verified data with its corresponding questionnaire before implementation by the data cleaning staff.

Output: Data cleaning program.

*ICTs***:** PCs, MS Word and Excel, DOS, Windows 2000/XP, CSPro, IMPS, C++, BASIC, and LAN are used

Head office – data processing, data cleaning:

Input: Edited and coded questionnaire, data (verified and structurally edited), Data cleaning program, PCs, MS Word and Excel, DOS, Windows 2000/XP, CSPro, IMPS, LAN, printers

Participants: Programmers, systems analysts, data cleaning staff, SMS, and editing and coding staff.

Process: Staff in the systems and programming division explains how the cleaning system works to the data cleaning staff responsible for the survey. The data cleaning staff receive the cleaning programs from the systems and programming division and the verified and structurally edited data from data entry operations staff. They correct all types of errors on the verified data they have received, based on the errors reported by the program. A print report containing the list of questionnaires with their IDs and corresponding errors is sent to the editing and coding staff. The editing and coding staffs collect the questionnaires in question, make the necessary corrections and send them back to the cleaning staff where they modify the verified data for each questionnaire, using the same data entry system as the data entry operators. They run the program again to check if it is free from errors. Finally they print out the counts for households and EAs and send them to the methodology service. While in this phase the data cleaning or editing and coding staffs may find unexpected errors that have not been removed by the program, edit specifications or both. They may also some errors identified by the program were not errors at all. In this case, both the SMS and systems and programming staff intercede and modify both the edit specifications and the program.

Output: Cleaned data, soft and hard copy of count of households and EAs.

*ICTs***:** PCs, MS Word and Excel, DOS, Windows 2000/XP, CSPro, IMPS, C++, BASIC, and LAN are used

Head office – methodology service:

Input: Soft and hard copy of count of households and EAs, sampling frame, PCs, MS Word and Excel, DOS, Windows 2000/XP, DBASE IV, and LAN

Participants: Methodology experts in the methodology department.

Process: The methodology service staff compare the counts of households and EAs against the code book. In some site areas some households might be missing or some EAs might even be closed. If there is a substantial difference in counts with no documented proof of missing households or EAs, it may mean that some questionnaires were not returned from the field, were not entered, or were not cleaned. The relevant branch offices and departments are then contacted to fill in the missing counts. Once the count is correct, weighting factors are prepared and sent to the systems and programming division to be used in tabulation.

Output: Weighting factor

*ICTs***:** PCs, MS Word and Excel, DOS, Windows 2000/XP, DBASE IV, and LAN are used

Head office – subject matter specialists (SMS):

Input: questionnaire, PCs, MS Word and Excel, Windows 2000/XP, and LAN

Participants: SMS

Process: Prepare tabulation plan to be used by the systems and programming division.

Output: Tabulation plan

*ICTs***:** PCs, MS Word and Excel, Windows 2000/XP, and LAN are used

Head office – data processing, systems and programming*:*

Input: Edited and coded questionnaire, cleaned data, tabulation plan, PCs, MS Word and Excel, DOS, Windows 2000/XP, CSPro, IMPS, SPSS, BASIC, and LAN

Participants: Programmers, systems analysts, and SMS

Process: The systems and programming division develops a program that reformats the cleaned data when necessary, do further cleaning, combine data sets and produce tables. The final data set is also converted to SPSS format, which is used for further analysis by users. On request the data set can also be changed to other data formats such as SAS, STATA, CSV, etc. It is then sent to ICT department to be archived

Output: counts of households and EAs, total household estimates, disaggregated statistical tables (by place, gender, age, etc), tables with estimates of some variables with standard error and confidence intervals.

*ICTs***:** PCs, MS Word and Excel, DOS, Windows 2000/XP, CSPro, IMPS, SPSS, BASIC, and LAN are used

Head office - methodology department:

Input: Soft copy of count of households and EAs, estimates of total households, and selected tables with estimates, standard error, and confidence interval, PCs, MS Word and Excel, Windows 2000/XP, and LAN

Participants: Methodology experts, programmers, system analysts.

Process: The methodology of service staff compares the counts of households and EAs with the code book. They also check the total household estimates with standard error and confidence intervals, sampling and non-sampling errors. If the estimates are unlikely, they contact the systems and programming division to check for any problems in the program

Output: Report of procedures and formulae used for estimation to produce the statistical tables, the number and size of households and the number of enumeration areas covered in the survey including the methodology.

ICTs: PCs, MS Word and Excel, Windows 2000/XP, and LAN are used

Head office - subject matter specialists (SMS):

Input: Disaggregated statistical tables (by place, gender, age, etc), PCs, MS Word and Excel, Windows 2000/XP, and LAN

Participants: SMS

Process: Prepare write up, further analysis using the past and current statistical tables.

Output: Report.

ICTs: PCs, laptops, MS Word and Excel, Windows 2000/XP, EURO Tracer, and LAN are used

Note: the SMS and methodology staff discuss over the results with the management before it is sent for printing of publication.

Head office - printing:

Input: Write-ups, printing machines, papers, PCs

Participants: Printing staff

Process: Publishing the write-ups using in-house printing machines.

Output: Publication (report).

ICTs: PCs, MS Word and Excel, Windows 2000/XP are used

Head office – ICT department:

Input: SPSS format of the final data set used for tabulation, write up, questionnaires, training manuals, code books, PCs, laptops, MS Word and Excel, Windows 2000/XP, Windows server operating systems, IHSN, Nestar, SPSS, CSPro, CDs, CD publishers, PDF makers, Arc/GIS, and LAN

Participants: ICT staff (responsible to archiving and dissemination)

Process: After confirmation that the reports are final or after publication, data and metadata including questionnaire, manuals, write-up, and statistical tables are archived in a server. They are also burned on CDs for dissemination on request. All

the information related to the survey is stored in the backup server. In addition to CDs information is also disseminated through the website.

Output: Archiving and backup of survey information, CDs, web links to information.

*ICTs***:** Data base and Web servers, PCs, laptops, MS Word and Excel, Windows 2000/XP, CDs, CD publishers, PDF makers, IHSN and Nesstar software for cataloguing surveys and data management, Arc/GIS software to display geo-referenced data on a map, and the website to disseminate information in addition to the CDs and publications are used

The ministry of Finance and Economic Development (MoFED):

Inputs: Data and information (report), PCs, laptops, MS Word and Excel, Windows 2000/XP, SPSS, STATA

Participants: MoFED Welfare Monitoring Unit (WMU) staff if the surveys are the case of WMS and HICES.

Process: Analysis of WMS, HICES, and data from other line ministries.

Output: Analytical report, progress report

*ICTs***:** PCs, laptops, MS Word and Excel, Windows 2000/XP, SPSS, STATA are used in analysis

6.4 Conclusion

In this chapter flow charts and description that are used to interpret how socio-economic data and information are generated through household surveys was discussed. The description and flow charts discussed are based on observations both at the head and branch offices. The observation during the pre-data collection stage includes questionnaire design, training of trainers, and training of enumerators. During data collection it includes field data collection and supervision. Observation during the post-data collection stage includes how data is captured, manually edited and coded, cleaned using computers, manipulated, used in tabulation. In addition in all stages the various roles of internal stakeholders and interactions, the different ICTs (hardware, software, and communication technologies) used were observed. This enables to understand some the effects on data quality if the data collection mode changes from PAPI to CAPI. It also enables to understand the effect on enumerators in the field as well as personnel working in the data entry, data cleaning, and data editing and coding sections. The next chapter gives a conclusion of the research and some recommendations based on the analysis made above.

CHAPTER SEVEN

CONCLUSION AND RECOMMENDATION

7.1 Conclusion

Key socio-economic indicators from household surveys mainly from the welfare monitoring surveys and household income consumption and expenditure surveys, population and housing census, and demographic and health surveys (DHS) are used by countries and international organizations to measure and evaluate progress in income poverty reduction, measure the extent and causes of poverty, and draw poverty line. They are also used to plan appropriate policies and interventions which are going to be implemented by the collaborative effort of concerned sector ministries and lower level administrative offices as well as regional, continental, and international organizations.

The assessment of country's progress with respect to human development such as the ranking of countries in the HDI and HPI, in the annual Human Development Reports prepared by the UNDP, is based on international data which in turn are obtained from national statistical offices and other data producing agencies. Other annual reports such as the WDR prepared by the World Bank heavily use data sets generated by NSOs. National statistical offices conduct quantitative surveys and other research and higher institutions conduct qualitative surveys to generate data and information.

Evidence is essential to make better decisions, monitor and evaluate anti-poverty programs. It is obtained from a number of sources. National statistical offices, among others, provide such evidences in the form of data and information. Their product is important inputs to better policies and strategies. The quality of data and information generated by data producing agencies therefore should be preserved and improved. Various ICTs are used in the different statistical processes in order to produce poverty-related raw data and information and

investigating how these data are produced and the relation and role of ICTs in improving the quality of these data and information is important.

In this research it was examined how preserving and improving the quality of data and information produced by data-producing agencies such as the national statistical offices play a vital role to reduce poverty. In order to do this the concept of poverty was first discussed from different points of views, namely the income approach and the capability approach. It was discussed how poverty is perceived and measured by different countries, organizations, and scholars.

It was found out that the different views draw on data from NSOs (from household and establishment surveys) and other data sources (such as central banks, customer services and administrative records from line ministries, and other research institutions, etc). In the case of Ethiopia, some of the data sets they obtain from CSA include the HICE, WMS, and DHS.

It was found out in this research that the development of effective polices, strategies and various anti-poverty programs rely on the quality of data and information obtained from data producing agencies. Improving data quality requires a coordinated and concerted effort by all stakeholders involved in the production of the data along with better and appropriate ICTs, and systems.

It was found out from the reports of various external stakeholders that there is some data quality weakness. The most common problems mentioned could be categorized as *accessibility* (in terms of availability), *serviceability* (in terms of lag), or *accuracy and reliability* (in terms of unit or item non-response rates). Proper and appropriate ICTs along with various information systems components (inputs, process, and outputs) that are used before data collection, during data collection and after data collection were discussed in relation to how they improve the quality of poverty-related data and information based on the DQAF.

It was indicated that based on the IMF's DQAF and with the use of appropriate ICTs, specific data quality dimensions that could be improved before data collection may include t*he prerequisites of quality, methodological soundness, and accuracy and reliability*. During the data collection process *accuracy and reliability, and serviceability dimensions* could be improved, and during the post-data collection process it was indicated that *accuracy and reliability, serviceability and accessibility* dimensions could be improved.

In this research it is recommended, as shown in the next section, that mobile and web-based data collection using portable PCs or PDAs (or their newest and enhanced versions) and internet respectively along with other appropriate ICTs and systems as a solution to address specific data quality problems. This should take in to account the expansion of the telecommunication and power infrastructures and improved quality of service. In addition to this comprehensive planning and careful assessment of the types of ICTs to be used, budget, training of experts, and maintaining staff turnover among other things is required. In the case of CSA, apart from the long-term cost saving, some of the data quality dimensions that could be improved may include:

Before data collection

Methodological soundness, and Accuracy and reliability (lessons and sharing of firsthand experience to have standard concepts and definitions in order to reduce interviewer errors)

During data collection

Accuracy and reliability, Serviceability (by reducing time)

During post-data collection

Accuracy and reliability (errors), *serviceability* (timeliness), *Serviceability* (timeliness)

In general, it was found out that regardless of the various views of the concepts of poverty, the research indicates that the use of appropriate ICTs in the various stages of the statistical process in NSOs may improve the quality of poverty-related data and information which are used as inputs in the process of monitoring and evaluation, analysis of poverty datum lines, development of PRSPs, anti-poverty programs, etc.

7.2 Recommendation

Below are some recommendations that may improve the capability of the CSA to generate and deliver quality data and information. They included suggestions about mobile data collection using portable computers, exploiting the potential of existing ICT infrastructure in data mining and the use of intranets to exchange experiences among internal stakeholders.

The recommendation on using the CAPI and Web-based data mode of data collection is based on the observed ICT infrastructure (mainly portable PCs) and availability of expertise in using these ICTs in relation to improving specific data quality dimensions.

7.2.1 Mobile Data Collection

NSOs use various ICTs at different stages in the statistical process to collect, process, store, analyse and disseminate data and information from household and establishment surveys. The use of portable computers in mobile data collection can improve certain quality dimensions, generated at the level NSOs, mainly those that result from interviewer or enumerator errors and non-response. Some of the data quality dimensions that can be improved by using the CAPI method may include:

- *Accuracy and reliability: D*ata quality dimension can be improved mainly through the reduction of item and unit non response rates,
- S*erviceability* (the timeliness element) data quality dimension can be improved by reducing the *time* between data collection and publication of survey results. This allows external main stakeholders such as MOFED and other line ministries to make timely analysis of data, and make informed decisions. It also allows various other external stakeholders such as the WB, UNDP, etc to compare global multidimensional poverty trends.
- The *respondent burden* could be minimized, depending on the nature of the survey.

- *Data quality* could also be enhanced by sequencing questions to the right household member (e.g. woman, child, holder, etc) or item, or by giving feedback in the form of error messages when outliers are entered or to items that are bounded with specific ranges, and other consistencies.

The selection and use of different portable PCs such as PDAs, or UMPCs in household surveys requires careful planning because:

- The technologies change at an increasing rate and each new product brings major changes. Some of these may include size and colour of the screen and fonts for visibility of electronic questionnaires, more processing power and speed according to whether software runs on the portable PC or is used in remote data capture only, the type of built-in communication technology for data transfer, memory capacity and other accessories such as batteries and their packs, covers, etc.
- Training of enumerators as well as trainers in the use of CAPI is different from that for PAPI.
- Data security issues should be addressed while data is being collected, stored and transferred from the field.
- A proper system including human resources that can address failures in hardware, software and communication problems should be in place.
- Field supervision using computer methods on the field needs to be devised as there is no questionnaire to be checked.
- The appropriateness of the CAPI mode of data collection using portable PCs for some surveys (e.g. fear of giving out information, which the respondent thinks is sensitive) needs to be well assessed for possible improvements to existing data quality in the PAPI mode.
- The roles and tasks of data entry staff, cleaning and other personnel that may be affected by this system need to be addressed.

In the developed world, where there is a well developed institutional and national ICT infrastructure and higher penetration of ICT usage, mobile data collection has already been used for a long time to collect most survey data from households and enterprise surveys. Some European and Latin American countries use Blaise software, named after Blaise Pascal and developed by the Statistics Bureau of Netherlands, to collect and process survey data. They use CAPI, CATI, CASI, and web-based modes of data collection with a laptop computer. In the developing world, mobile data collection is used mainly in the health sectors. In the case of Ethiopia, CSA uses CAPI mode for the consumer price indices survey (as discussed in section 5.5) and the Ministry of Health (MoH) used CAPI mode in 2007 for the malaria indicator survey. CSA began mobile data collection using CAPI mode using PDAs for the consumer price index survey at selected market places in 2007. Microsoft Excel is used to develop the form that is used to capture the data. It was reported that the time of delivery, an element of the *serviceability* data quality dimension, has improved as a result. The CAPI application of the MoH however reported some logistics, human resources, hardware, and software difficulties (as discussed in section 4.4).

The CSA has been collecting data using PAPI method and the data is captured and processed at the head office using CSPro software, public domain software, developed by the international program centre of the United States Census Bureau, Macro International, and Serpro, SA. IMPS software, a DOS-based tool, is also used in the processing of data. Like Blaise CSPro it was developed specifically for statistical offices for surveys and censuses. CSPro has been used to design a questionnaire for the CAPI mode of data collection instrument to collect data using laptops. The CAPI mode can also be used in PDAs in mobile data collection. One limitation however is that at the time of writing it cannot be used in smart phones for data collection. The CSA has been using this software for a long time in data entry, cleaning, and processing during the post-data collection stage.

With the existing infrastructure and experience of using the CSA's CSPro software can apply the CAPI mode of data collection to more surveys. The table below shows the changes that may take place in the current system, and where possible the data quality improvements that could be obtained as a result of the CAPI mode of data collection using portable PCs. It is

followed by a flow chart that shows the data flow in the CAPI mode of data collection and its explanation.

	Current process	With mobile data collection	Data quality
	Pre data collection		
1	Questionnaire design	Same	*Methodological soundness, and Accuracy and reliability* (lessons and sharing of firsthand experience to have standard concepts and definitions in order to reduce interviewer errors)
2	Printing questionnaire	No need Soft copy using CAPI on portable PCs	
3	Printing manuals and codebooks	No need. Soft copy is used on portable PCs (PDF or word files)	
4	Training of trainers	Do. Does include training on the use of CAPI, computer literacy, and troubleshooting.	
5	Dispatching trainers	Do.	
6	Dispatching questionnaires and other survey documents	No questionnaire. Soft copy on portable PC	
7	Manuals	No manuals	
8	Training of:		
9	Training of enumerators	The same, but includes training in the use of CAPI, computer literacy, and troubleshooting.	
10	Training of supervisors	The same, but includes training in the use of CAPI, computer literacy, and troubleshooting. Addition supervisory methods to monitor the CAPI mode of data collection.	
	During Data Collection		
11	Data collection	CAPI	*Accuracy and reliability, Serviceability* (by reducing time)
12	Field supervision	CAPI	
	Post Data Collection		
13	Collecting questionnaires/data transfer	No questionnaire. Data is transferred using built in GPRS or CMDA, e-mail to head office or branch office, or using a USB cable or Bluetooth or wireless connectivity to branch office PCs	*Accuracy and reliability* (errors), *serviceability* (timeliness), *Serviceability* (timeliness)
14	Branch office editing	Using PCs	
15	Transporting questionnaires	No transportation	
16	Documentation	No hard copy documentation	
17	Head office editing and coding	Using PCs and laptops	
18	Data entry	No entry	
19	Data cleaning	Minimal	
20	Data reformatting	Same	
21	Tabulation	Same	
22	Writing survey report	Same	
23	Printing survey report	Same	

Table 7-1 Possible Data Quality Improvements with Mobile Data collection

(Source: Author's interpretation)

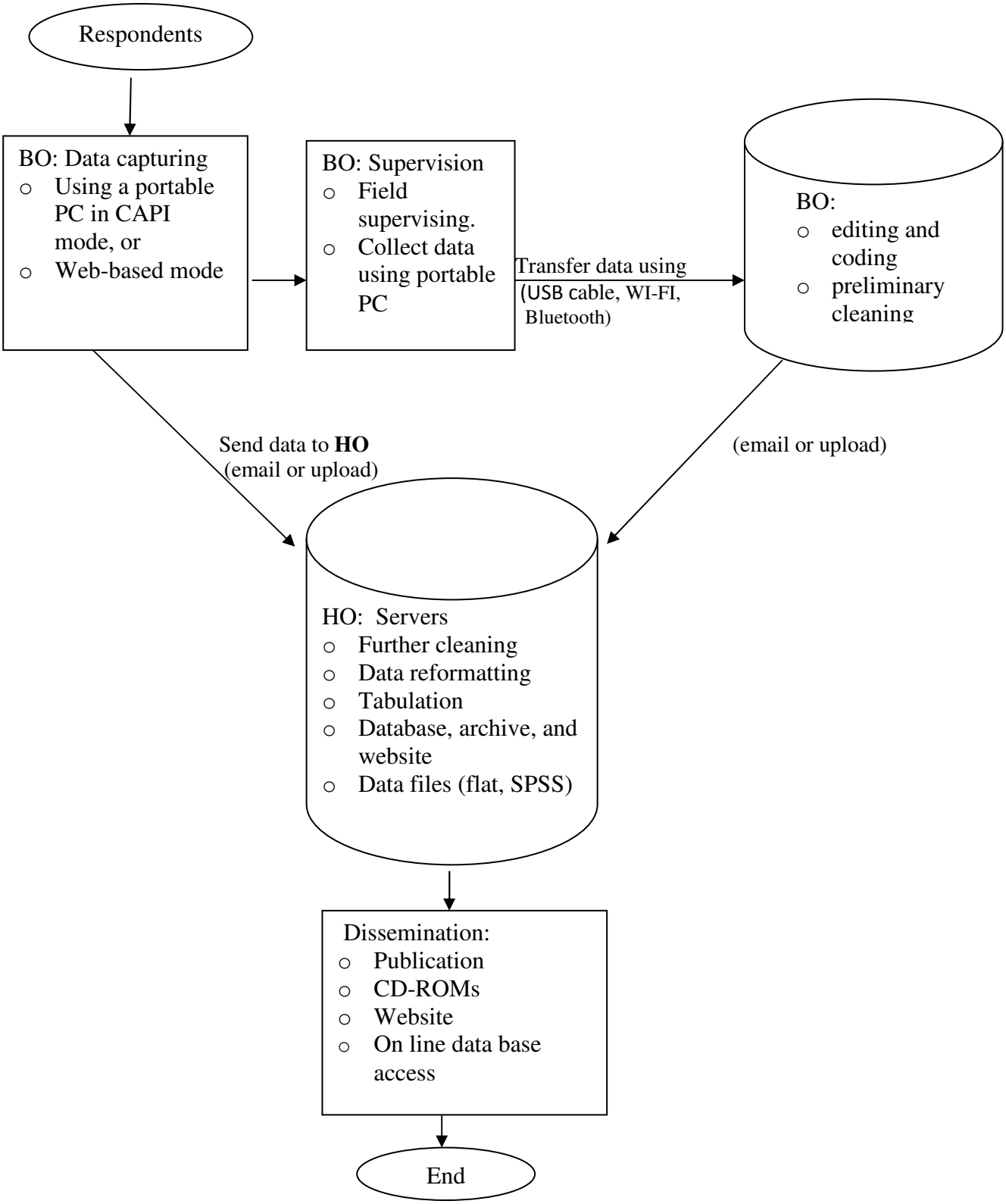


Figure 7-1 Data Flow under mobile data collection

(Source: Author's illustration)

Note: *BO*-branch office, and *HO*-head office

The general data flow involves the use of CAPI with the help of a PDA or other portable PC. The portable PCs must be compatible with the CAPI application. The CAPI application, electronic version (PDF or Word files) of all survey documents that are necessary for the training and data collection such as questionnaires, training manuals, codebooks, EA maps, etc. are copied into the portable PCs. The CAPI application and the portable PCs with all their accessories should be tested and checked well before training begins, to minimize unnecessary time spent troubleshooting the software and hardware systems. Sufficient time should be allowed for the pilot testing of the CAPI application in different branch offices. This reduces the risk of having bugs and defective hardware and provides useful feedback. When the system is running, training of trainers can be conducted in a similar fashion as for PAPI, including the use of the portable PCs, the CAPI application and basic troubleshooting with pilot testing. When this training is completed, the body concerned makes sure that all portable PCs are non-defective, have the necessary backup and have CAPI and other survey information loaded. Enumerators use these portable PCs during the training in the branch office in a similar fashion as with PAPI and conduct practical interviews in the field as a pilot test. Afterwards the pilot feedback at national level is collected and hardware, programming, connectivity, and other problems get fixed for smooth data collection.

Field Data Collection

First enumerators capture data in face-to face interviews from households using CAPI mode with the help of a portable PC. First data is entered in to the electronic questionnaire using input devices such as styles, built-in key board, or external key board. While data is entered different types of errors are also corrected. The data can be saved on the portable PC (or SD and memory sticks if necessary). The collected data then can be sent to the branch office or directly to the head office. A web-based self-administered data can also be collected for selected surveys such as the quarterly industry survey.

Branch Offices

The captured data set can be collected from the enumerators' portable PCs in a number of ways: supervisors can collect the data from the field, or enumerators can bring their devices to the branch office and to be linked to desktop PCs. In the case of data collected using the web-based self-administered mode, data can be sent directly to Head Office for further

processing. This can be an e-mail attachment or directly on the database that is accessible to the respondent.

Supervisors can collect data from enumerators from the field (or the enumerators may bring the data to the branch office) after checking for consistencies, ranges, listings, etc in the same way as standard, but with the help of different computer programs installed on laptops or other portable PCs. This has the advantage of correcting problems by re-interviewing or re-measurement to improve data quality before leaving the enumeration area. Supervisors then take the data to the branch office to transfer the data (using USB cables, wireless or Bluetooth) to the desktop PCs for first stage or final computer editing and cleaning. Finally the branch offices send the cleaned data to Head Office. If there is no need of editing and cleaning at the branch office and there is a reliable telecommunication infrastructure, enumerators can send the data directly to Head Office.

Head Office

At the head office depending on the need of further processing for different surveys, additional data cleaning and editing can be done. When all the processing such as further data cleaning, data reformatting, tabulation, and archiving is completed all survey information is uploaded to the website to be accessed by stakeholders. Survey information can also be disseminated and communicated in various other ways such as publication, workshops, and CD-ROMs.

7.2.2 Potential of Using Existing ICT Infrastructure to Other Purposes

At the time of this writing CSA is developing a centralized relational database management system for most of its surveys using MS-SQL 2008. The objective is to be able to provide stakeholder a time series data for online analysis, and a complete system for data processing components such as data entry, cleaning, and tabulation. CSA is conducting various socio-economic surveys in an integrated method, i.e., WMS, HICES, agricultural, and other surveys are conducted in the same primary sampling units[212] . This means that if the reporting level for all the surveys in the future is at the district level as per the recommendation of some

[212] Primary sampling units (PSU) are units within the lowest administrative level that are only used by CSA for the purpose of data collection. Within one lowest administrative level there might be more than three PSUs. In agricultural surveys for example, within one PSU there are about 250 to 300 households from which about 30 households are selected for interview.

stakeholders (including the WB), one can have a full picture of the district enabling planners to take action and follow up progress in their respective district. More over it is possible to use data mining technologies which are integrated in MS-SQL server 2008 software to further analysis and to find out interesting relationships between time series surveys and across surveys.

Intranet, as a collaboration tool, its implementation and use enables employees at different levels to interact horizontally to share ideas, feedbacks, solutions, and experiences to enhance organizational efficiency. Sharing of information during training of trainers at the head office and training of enumerators in the branch office takes place. The information is like a feedback based on what trainers of enumerators, supervisors, technicians, and enumerators faced on the real job on the ground. They are not archived in a structured digital form to be accessible via CSA's website to its internal stakeholders. This has advantage in improving *data quality* as data collection is one of the most important steps requiring great effort and techniques to acquire the right information from respondents. It enables internal stakeholders of CSA to improve the survey operation (or standardize it) if these feedbacks are shared on the website or intranet (when developed). With the current available ICT infrastructure of CSA, it is possible to develop intranet to be used within the head office. It is also possible to include external stakeholders to use CSA's internal resources by assigning different levels of access.

BIBLIOGRAPHY

Alkire, S. and Deneulin, S. 2009. The Human Development and Capability Approach *In* Deneulin, S. and *Shahani L. (eds.), An introduction to the human development and capability approach : freedom and agency.* UK and USA: Earthscan.

Artemeva, N. 2006. Communities of Practice: Learning, Meaning, and Identity. Reviewed in: *Technical Communication Quarterly. 15*(4), 505-507.

Asian Development Bank (ADB) 2004, Poverty *Profile of the People's Republic of China*, Manila.

Asian Development Bank (ADB) 2005, *Country Strategy and Program Update (2006-2008)*, Manila. Retrieved August 27, 2009, from http://www.adb.org/Documents/CSPs/PRC/2005/default.asp.

Bahra, N. 2001, *Competitive Knowledge Management*, Palgrave: New York.

Bjorkqvist, Sven et.al 2002. Experiences with security in statistics Finland. Statistical Journal of the UN Economic Commission for Europe, 2002, 19(3), pp. 211–218.

Boisot, M. 1998. *Knowledge assets – securing competitive advantage in the information economy.* Oxford: New York.

Bresciani, F and Valdes, A. (eds.) 2007. *Beyond food production: the role of agriculture in poverty reduction.* Rome: FAO.

Carson, Carol S 2001. *Toward a Framework for Assessing Data Quality, IMF Working Paper,* WP/01/25. Washington: International Monetary Fund.

Central Statistics Agency of Ethiopia 2004a. *Directive Issued To Establish Procedures For Accessing Raw Data To Users*, Directive No. 1 /2004. Retrieved June 1, 2009 from http://www.csa.gov.et/index.php?option=com_content&view=article&id=124 &Itemid=82.

Central Statistical Agency of Ethiopia (CSA) 2004b. *Welfare Monitoring Survey:*

analytical report, Retrieved Feb 12, 2010, from http://www.csa.gov.et/index.php?option=com_wrapper&view=wrapper&Itemid=185.

Central Statistical Agency of Ethiopia (CSA) 2005a. *Central Statistics Agency Establishment Proclamation* No. 442/ 2005. Retrieved June 1, 2009, from http://www.csa.gov.et/index.php?option=com_content&view=article&id=124&Itemid=82 .

Central Statistical Agency of Ethiopia (CSA) 2005b. *Strategic plan document* (Amharic version). Retrieved June 1, 2009, from http://www.csa.gov.et.

Central Statistical Agency of Ethiopia (CSA) 2007a. 2004/05 Household Income, Consumption and Expenditure Survey (HICE):statistical report Volume II. *analytical report.* Retrieved Oct 12, 2010, from http://www.csa.gov.et/index.php?option=com_wrapper&view=wrapper&Itemid=547.

Central Statistical Agency of Ethiopia (CSA) 2007b, *Summary and Statistical Report of The 2007 Population and Housing Census: population size by age and sex*, UNFPA, Addis Ababa, Retrieved Feb 12, 2010, from http://www.csa.gov.et/pdf/Cen2007_firstdraft.pdf.

Central Statistical Agency of Ethiopia (CSA) 2008, *Agricultural sample survey 2007/2008 (2000 E.C.), Volume IV, Report on Land Utilization (private peasant holdings, Meher season)*, Retrieved Feb 12, 2010, from http://www.csa.gov.et/index.php?option=com_wrapper&view=wrapper&Itemid=170.

Ethiopian Central Statistical Agency (CSA) 2009. *National Statistical Development Strategy (NSDS).* Retrieved May 13, 2010, from http://www.csa.gov.et.

Ethiopian Central Statistical Agency (CSA) 2009. *National Statistical Development Strategy (NSDS).* Retrieved May 13, 2010, from http://www.csa.gov.et.

Checkland, P. and Holwell, S. 1998. Information, Systems and Information systems: making sense of the field. Chichester: Wiley.

Chen, Shaohua and Martin Ravallion 2004. *How have the world's poorest fared since the early 1980s?* Policy Research Working Paper 3341. Washington DC: World Bank.

Chen, Shaohua and Martin Ravallion 2008. *The developing world is poorer than we thought, but no less successful in the fight against poverty*. Policy Research Working Paper 4703. Washington DC: World Bank.

Cheru, F. 2006. Building and Supporting PRSPs in Africa: what has worked well so far? What needs changing? Third World Quarterly, 27(2), pp 355 – 376.

Dixon and Macarov 1998, *Poverty: a Persistent Global Reality*. Routledge, 11 New Fetter Lane, London EC4P 4EE.

David Loshin 2001, *Enterprise Knowledge Management: the data quality approach*. Academic press, California, USA.

Ellis, F. and Woldehanna, T. 2005, *Ethiopia Participatory Poverty Assessment 2004-05*, Addis Ababa: Ministry of Finance and Economic Development (MoFED), October.

EUROSTAT 2009. *ESS handbook for quality reports.* Retrieved 12, February 2010 from http://epp.eurostat.ec.europa.eu/portal/page/portal/product_results/search_results?mo=containsall&ms=ESS+Handbook+for+Quality+Reports&saa=&p_action=SUBMIT&l=us&co=equal&ci=,&po=equal&pi.

European Commission. 2010. *Combating poverty and social exclusion: a statistical portrait of the European Union 2010***.** Luxembourg: Publications Office of the European Union.

James M. Galliher, Thomas V. Stewart, Paramod K. Pathak, James J. Werner, L. Miriam Dickinson, & John M. Hickner 2008. Data Collection Outcomes Comparing Paper Forms With PDA Forms in an Office-Based Patient Survey. *Annals of Family Medicine* 6:154-160.

Ghauri, P. and Grønhaug, K. 2005. *Research Methods in Business Studies: a practical guide*. Harlow, England : Financial Times Prentice Hall.

Hulme, D. 2009. *The millennium development goals (MDGs): a short history of the world's biggest promise*. BWPI Working Paper 100. University of Manchester:BWPI. Retrieved April 10, 2010, from http://www.manchester.ac.uk/bwpi..

ITU 2009. *Measuring the information society: the ICT development index*. Switzerland, Geneva.

JM Carroll and PA Swatman 2000. Structured-case: a methodological framework for building theory in information systems research. *European Journal of Information Systems.* 9: 235–242.

Jodi et al. 2007. Use of Handheld Computers with Global Positioning Systems for Probability Sampling and Data Entry in Household Surveys. *American Journal of Tropical Medicine and Hygiene*, 77(2), 2007, pp. 393-399.

Keith Devlin 1999. *Infosense: turning information into knowledge*. New York: Freeman.

Kirubel Tadesse 2009**.** *Capital. Zero Import tax for wireless CDMA.* Local News: August 2009. Retrieved September 15, 2009, from http://www.capitalethiopia.com/archive/2009/August/week3/local_news.htm#7.

Klugman, J. (ed.) 2002, *Source Book For Poverty Reduction Strategies (Vol. 1 of 2): core techniques and cross-cutting issues*. Washington, DC. World Bank.

Kunzler, Uwe 2002. Electronic data reporting (EDR), metadata, standards and the European statistical system (ESS). Statistical Journal of the UN Economic Commission for Europe, 2002, 19(3), pp. 119-130.

Leedy, Paul D., Ormrod, Jeanne Ellis 2005. *Practical research: planning and design*. Upper Saddle River, N.J. : Pearson/Merrill Prentice Hall.

Marshal, C., Rossman, G.B. 1999. *Designing Qualitative Research*. SAGE, UK.

Ministry of Capacity Building (MoCB) 2006. *The National ICT for development five years action plan for Ethiopia (2006-2010).* Retrieved June 1, 2009 from http://www.mocb.gov.et, or from the Ethiopian ICT development agency (EICTDA) http://www.eictda.gov.et/?q=node/15.

Ellis, F. & Tassew, W. 2005. *Ethiopia participatory poverty assessment 2004–05*. Addis Ababa: Ministry of Finance and Economic Development (MoFED).

Ministry of Finance an Economic Development (MoFED) 2006, *A Plan for Accelerated and Sustained Development to End Poverty (PASDEP) (2005/06-2009/10), Volume I.* Retrieved May 12, 2009 from http://www.mofaed.gov.et.

Ministry of Finance and Economic Development (MoFED) 2007. *Ethiopia Building on Progress: A Plan for Accelerated and Sustained Development to End Poverty (PASDEP), Annual Progress Report 2006/07*, Retrieved August 16, 2009 from

http://www.mofed.gov.et/images/news/PasdepAnnualProgressReport2006-2007.pdf.

Ministry of Health (MoH) 2008. *Ethiopia malaria indicator survey 2007.* Retrieved on June 1, 2010 from http://www.malariasurveys.org/surveys.cfm?country=Ethiopia2007#5002.

Morse, Stephen 2004, *Indices and indicators in development: an unhealthy obsession with numbers?* London: Earthscan.

National Bank of Ethiopia 2006/07, *Overall Economic Performance, Annual Report: National Bank Report 2006* – 2007. Retrieved June 10, 2009 from http://www.nbe.gov.et/publications/annualreport.htm.

National Bank of Ethiopia 2007/08, *Overall Economic Performance, Annual Report: National Bank Report 2006* – 2007. Retrieved June 10, 2009 from http://www.nbe.gov.et/publications/annualreport.htm.

Neas, Kathy; Molina, Jubal; Hardegree, Jason; Gerling, Michael 2006. *Using Personal Digital Assistants for the 2004 Cotton Objective Yield Survey*. United States Department of Agriculture, National Agricultural Statistics Service. RDD research report number RDD-06-01. Retrieved March 10, 2010 from http://www.nass.usda.gov/Education_and_Outreach/Reseach_Reports/index.asp.

Nonaka I (ed). 2005, *Knowledge Management: Critical Perspectives on Business and Management,* London: Routledge.

Nutley, S.M., Isabel Walter and Huw T.O. 2007, *Using evidence: how research can inform public services,* Bristol: Policy Press.

OECD 2005. Joint european commission – OECD workshop on international development of business and consumer tendency surveys: Part 1 – Relationship between response rates and data collection methods. Retrieved Sept 10, 2010 from http://www.oecd.org/dataoecd/55/40/35558806.pdf

OECD 2006. OECD workshop on business and consumer tendency surveys: Internet Business Tendency Surveys. Retrieved Sept 10, 2010 from http://www.oecd.org/dataoecd/19/10/37325653.pdf.

Orlikowski, Wanda J *and* Baroudi, Jack J. 1991. Studying Information

Technology in Organizations: Research Approaches and Assumptions. *Information Systems Research. 2(1): 1-28.*

Park, A., and Wang, S. 2001, China's poverty statistics. *China Economic Review, 12(4):* 384-398.

Ravallion, M. 1998. *Poverty Lines in Theory and Practice, Living Standards Measurement Study*; Working Paper 133, World Bank, Washington DC.

Road *Map towards the Implementation of the United Nations Millennium Declaration*: Report of the Secretary-General, A/56/326, Sept. 2001, UN.

Robson, Colin 1993. *Real World Research: a resource for social scientists and practitioner-researchers*. Oxford: Blackwell publishers.

Rooney D, Hearn G, Ninan A (eds) 2005. *Handbook on the Knowledge Economy* Cheltenham, UK: Edward Elgar.

Roos, Marko 2002. Methods of Internet data collection and implications for recruiting respondents.Statistical Journal of the UN Economic Commission for Europe, 2002, 19(3), pp. 175–186.

Schräpler, J.-P., Schupp, J., & Wagner, G. G. 2006. Changing from PAPI to CAPI: a longitudinal study of mode-effects based on an experimental design (Discussion papers Vol. 593). *German Institute for Economic Research*. Berlin: DIW Berlin. http://www.diw.de.

Sen, A.K. 1999. Development as Freedom. New York: Knopf Press.

Spicker P, Leguizamon S & Gordon D (eds). 2007. *Poverty: An International Glossary. 2nd edition*. UK and New York: Zed Books Ltd.

Staire, Ralph M. & Reynolds, George W. 2001. *Principles of Information Systems: A managerial approach*. Boston, Mass.: Course technology.

Staire, Ralph M. & Reynolds, George W. 2006. *Principles of Information Systems: A managerial approach*. Boston, Mass.: Course technology.

The World Bank 1997. *Taking action to Reduce Poverty in Sub-Saharan Africa*, Development in Practice series, Washington, DC: World Bank.

The World Bank 2002. *Information communication Technologies*: A World Bank Group Strategy. Washington, DC: World Bank.

The World Bank 2007. *World development indicators 2007*. Washington: World Bank.

Thomas H. Davenport and Laurence Prusak 1997, *Information Ecology: Mastering Information and Knowledge Environment*, New York: Oxford University Press.

United Nations (UN) 2008. *The Millennium Development Goals Report*. New York: UN Department of Economic and Social Affairs.

United Nations (UN) 2009. *The Millennium Development Goals Report*. New York: UN Department of Economic and Social Affairs.

UNDP 2007. *Human Development Report 2007/2008 :fighting climate change : human solidarity in a divided world*. New York : Palgrave Macmillan.

UNDP 2009. *Human Development Report 2007/2008 overcoming barriers : barriers :human mobility and development*. New York : Palgrave Macmillan.

U.S. Census Bureau 2006, *Definition of data quality: Census Bureau Principle*, Version 1.3, June 14. Retrieved July 2, 2009 from http://www.census.gov/quality.

Vivoda and Eby 2006. Using Personal Digital Assistants (PDAs) for the Collection of Safety Belt Use Data in the Field. *Behaviour Research Methods.* 38 (1):158-164.

Walsham G 1995. Interpretive case studies in IS research: nature and method. *European Journal of Information Systems. 4(2):* 74–81.

Walsham G 2006. Doing interpretive research. *European Journal of Information Systems. 15(3):* 320–330.

APPENDIX

1. Appendix A

The full list of the millennium development goals (MDGs) with goals, targets, and indicators for monitoring progress is listed in the following table:

Millennium Development Goals (MDGs)	
Goals and Targets (from the Millennium Declaration)	**Indicators for monitoring progress**
Goal 1: Eradicate extreme poverty and hunger	
Target 1.A: Halve, between 1990 and 2015, the proportion of people whose income is less than one dollar a day	○ Proportion of population below $1 (PPP) per day[213] ○ Poverty gap ratio ○ Share of poorest quintile in national consumption
Target 1.B: Achieve full and productive employment and decent work for all, including women and young people	○ Growth rate of GDP per person employed ○ Employment-to-population ratio ○ Proportion of employed people living below $1 (PPP) per day ○ Proportion of own-account and contributing family workers in total employment
Target 1.C: Halve, between 1990 and 2015, the proportion of people who suffer from hunger	○ Prevalence of underweight children under-five years of age ○ Proportion of population below minimum level of dietary energy consumption
Goal 2: Achieve universal primary education	
Target 2.A: Ensure that, by 2015, children everywhere, boys and girls alike, will be able to complete a full course of primary schooling	• Net enrolment ratio in primary education • Proportion of pupils starting grade 1 who reach last grade of primary • Literacy rate of 15-24 year-olds, women and men
Goal 3: Promote gender equality and empower women	

[213] For monitoring country poverty trends, indicators based on national poverty lines should be used, where available.

[213] The actual proportion of people living in slums is measured by a proxy, represented by the urban population living in households with at least one of the four characteristics: (a) lack of access to improved water supply; (b) lack of access to improved sanitation; (c) overcrowding (3 or more persons per room); and (d) dwellings made of non-durable material.

Target 3.A: Eliminate gender disparity in primary and secondary education, preferably by 2005, and in all levels of education no later than 2015	○ Ratios of girls to boys in primary, secondary and tertiary education ○ Share of women in wage employment in the non-agricultural sector ○ Proportion of seats held by women in national parliament
Goal 4: Reduce child mortality	
Target 4.A: Reduce by two-thirds, between 1990 and 2015, the under-five mortality rate	○ Under-five mortality rate ○ Infant mortality rate ○ Proportion of 1 year-old children immunised against measles
Goal 5: Improve maternal health	
Target 5.A: Reduce by three quarters, between 1990 and 2015, the maternal mortality ratio	○ Maternal mortality ratio ○ Proportion of births attended by skilled health personnel
Target 5.B: Achieve, by 2015, universal access to reproductive health	○ Contraceptive prevalence rate ○ Adolescent birth rate ○ Antenatal care coverage (at least one visit and at least four visits) ○ Unmet need for family planning
Goal 6: Combat HIV/AIDS, malaria and other diseases	
Target 6.A: Have halted by 2015 and begun to reverse the spread of HIV/AIDS	○ HIV prevalence among population aged 15-24 years ○ Condom use at last high-risk sex ○ Proportion of population aged 15-24 years with comprehensive correct knowledge of HIV/AIDS ○ Ratio of school attendance of orphans to school attendance of non-orphans aged 10-14 years
Target 6.B: Achieve, by 2010, universal access to treatment for HIV/AIDS for all those who need it	○ Proportion of population with advanced HIV infection with access to antiretroviral drugs
Target 6.C: Have halted by 2015 and begun to reverse the incidence of malaria and other major diseases	○ Incidence and death rates associated with malaria ○ Proportion of children under 5 sleeping under insecticide-treated bed nets ○ Proportion of children under 5 with fever who are treated with appropriate anti-malarial drugs ○ Incidence, prevalence and death rates associated with tuberculosis ○ Proportion of tuberculosis cases detected and cured under directly observed treatment short course
Goal 7: Ensure environmental sustainability	

Target 7.A: Integrate the principles of sustainable development into country policies and programmes and reverse the loss of environmental resources Target 7.B: Reduce biodiversity loss, achieving, by 2010, a significant reduction in the rate of loss	5) Proportion of land area covered by forest 6) CO2 emissions, total, per capita and per $1 GDP (PPP) 7) Consumption of ozone-depleting substances 8) Proportion of fish stocks within safe biological limits 9) Proportion of total water resources used 10) Proportion of terrestrial and marine areas protected 11) Proportion of species threatened with extinction
Target 7.C: Halve, by 2015, the proportion of people without sustainable access to safe drinking water and basic sanitation	12) Proportion of population using an improved drinking water source 13) Proportion of population using an improved sanitation facility
Target 7.D: By 2020, to have achieved a significant improvement in the lives of at least 100 million slum dwellers	14) Proportion of urban population living in slums[214]
Goal 8: Develop a global partnership for development	
Target 8.A: Develop further an open, rule-based, predictable, non-discriminatory trading and financial system Includes a commitment to good governance, development and poverty reduction – both nationally and internationally Target 8.B: Address the special needs of the least developed countries	*Some of the indicators listed below are monitored separately for the least developed countries (LDCs), Africa, landlocked developing countries and small island developing States.* <u>Official development assistance (ODA)</u> o Net ODA, total and to the least developed countries, as percentage of OECD/DAC donors' gross national income o Proportion of total bilateral, sector-allocable ODA of OECD/DAC donors to basic social services (basic education, primary health care, nutrition, safe water and sanitation) o Proportion of bilateral official development assistance of OECD/DAC donors that is untied o ODA received in landlocked developing countries

[214] The actual proportion of people living in slums is measured by a proxy, represented by the urban population living in households with at least one of the four characteristics: (a) lack of access to improved water supply; (b) lack of access to improved sanitation; (c) overcrowding (3 or more persons per room); and (d) dwellings made of non-durable material

Includes: tariff and quota free access for the least developed countries' exports; enhanced programme of debt relief for heavily indebted poor countries (HIPC) and cancellation of official bilateral debt; and more generous ODA for countries committed to poverty reduction Target 8.C: Address the special needs of landlocked developing countries and small island developing States (through the Programme of Action for the Sustainable Development of Small Island Developing States and the outcome of the twenty-second special session of the General Assembly) Target 8.D: Deal comprehensively with the debt problems of developing countries through national and international measures in order to make debt sustainable in the long term	as a proportion of their gross national incomes ○ ODA received in small island developing States as a proportion of their gross national incomes <u>Market access</u> ○ Proportion of total developed country imports (by value and excluding arms) from developing countries and least developed countries, admitted free of duty ○ Average tariffs imposed by developed countries on agricultural products and textiles and clothing from developing countries ○ Agricultural support estimate for OECD countries as a percentage of their gross domestic product ○ Proportion of ODA provided to help build trade capacity <u>Debt sustainability</u> ○ Total number of countries that have reached their HIPC decision points and number that have reached their HIPC completion points (cumulative) ○ Debt relief committed under HIPC and MDRI Initiatives ○ Debt service as a percentage of exports of goods and services
Target 8.E: In cooperation with pharmaceutical companies, provide access to affordable essential drugs in developing countries	○ Proportion of population with access to affordable essential drugs on a sustainable basis

Target 8.F: In cooperation with the private sector, make available the benefits of new technologies, especially information and communications	○ Telephone lines per 100 population ○ Cellular subscribers per 100 population ○ Internet users per 100 population

Table 2.2a Goals, targets and indicators of the MDG

Source: The official United Nation's site for the MDGs indicators,

http://mdgs.un.org/unsd/mdg/Host.aspx?Content=Indicators/OfficialList.htm

2. Appendix B

The full list of the international telecommunications union (ITU) indicators are listed in the following table:

I61 Population	I153CO$ Mobile cellular - price of 3-minute local call (off-peak - US$)
I6111 Population – Urban population (%)	I153O Price of a 3-minute fixed telephone local call (off-peak rate)
I612 Population of largest city	I153O$ Price of a 3-minute fixed telephone local call (off-peak rate - US$)
I62 Households	I271 Mobile cellular telephone subscribers - (Post-paid + Pre-paid)
I63 Gross domestic product (GDP) - local currency	I2712 Mobile cellular telephone subscribers (Digital)
I63$ Gross domestic product (GDP) (US$)	I271P Mobile cellular telephone subscribers – prepaid subscribers
I64 Gross Fixed Capital Formation (GFCF) - local currency	I271POP Coverage of mobile cellular network (population, in %)
I64$ Gross Fixed Capital Formation (GFCF) (US$)	I51 Staff (Total full-time telecommunications staff)
I652 Average annual exchange rate per US$	I51F Staff (Female telecommunication staff)
I66_95 Consumer price index (1995=100)	I51W Mobile communications staff
I67700001 Telecommunication equipment (Export) (US$)	I71 Revenue from fixed telephone service
I67700002 Telecommunication equipment (Import) (US$)	I71$ Revenue from fixed telephone service (US$)
I1111 % of households with a telephone	I741 Revenue from mobile communication
I1112 Public payphones	I741$ Revenue from mobile communication (US$)
I112 Main (fixed) telephone lines in operation	I75 Total revenue from all telecommunication services
I1121 Main (fixed) telephone lines in largest city	I75$ Total revenue from all telecommunication services (US$)
I112t Total telephone subscribers (fixed + mobile)	I81 Total annual investment in telecom
I114 % automatic main lines	I81$ Total annual investment in telecom (US$)
I1142 % digital main lines	I83 Fixed telephone service investment
I116 % residential main lines	i83$ Fixed telephone service investment (US$)
I1162 % of main lines in urban areas	I841m Mobile communication investment
I117 Total capacity of local public switching exchanges	I841m$ Mobile communication investment (US$)
I1191 International telephone circuits	I91 Main (fixed) telephone lines per 100 inhabitants
I123 Waiting list for main (fixed) lines	I911 Mobile cellular telephone subscribers per 100 inhabitants

I1311C Number of local (fixed) telephone (calls)	I9111 Total telephone subscribers (fixed + mobile) per 100 inhabitants
I1311M Number of local (fixed) telephone (minutes)	I98 Public payphones per 1000 inhabitants
I1312C Number of national (fixed) long distance telephone (calls)	I955 Radio sets
I1312M Number of national (fixed) long distance telephone (minutes)	I955H Radio equipped households
I131C Total national (fixed) telephone traffic (calls)	I955H% % of households with a radio
I131M Total national (fixed) telephone traffic (minutes)	I965 Television receivers
I132C International outgoing fixed telephone traffic (calls)	I9651 Television receivers per 100 inhabitants
I132CI International incoming fixed telephone traffic (calls)	I965C Cable television subscribers
I132M International outgoing fixed telephone traffic (minutes)	I965H Television equipped households
I132MI International incoming fixed telephone traffic (minutes)	I965H% % of households with a television
I141 % of telephone faults cleared by next working day	I965S Home satellite antennas
I143 Faults per 100 main (fixed) lines per year	I28 ISDN subscribers
I151 Residential telephone connection charge	I28C ISDN Channels
I151$ Residential telephone connection charge (US$)	I4212 Estimated Internet users
I151B Business telephone connection charge	I4213 Total (fixed) Internet subscribers
I151B$ Business telephone connection charge (US$)	I4213CAB Cable modem Internet subscribers
I151C Mobile cellular connection charge	I4213D Dial-up Internet subscribers
I151C$ Mobile cellular connection charge (US$)	I4213DSL DSL Internet subscribers
I152 Residential monthly telephone subscription	I4213HP % of homes with Internet
I152$ Residential monthly telephone subscription (US$)	I4213TFB Total fixed broadband Internet subscribers
I152B Business telephone monthly subscription	I4214 International Internet Bandwidth (Mbps)
I152B$ Business telephone monthly subscription (US$)	I422 Personal computers
I152C Mobile cellular monthly subscription	I422HP % of homes with a Personal Computer
I152C$ Mobile cellular monthly subscription (US$)	I981 Personal computers per 100 inhabitants
I153 Price of a 3-minute fixed telephone local call (peak rate)	I99 Internet users per 100 inhabitants
I153$ Price of a 3-minute fixed telephone local call (peak rate - US$)	I992 Total fixed broadband Internet subscribers per 100 inhabitants
I153C Mobile cellular - price of 3-minute local call (peak)	I993 Total (fixed) Internet subscribers per 100 inhabitants
I153C$ Mobile cellular - price of 3-minute local	I994 International Internet Bandwidth per

call (peak - US$)	inhabitant (bit/s)
I153CO Mobile cellular - price of 3-minute local call (off-peak)	

Table 2.2b ICT indicators

Source: ITU

3. Appendix C

The table below indicates IMF's data quality assessment framework.

Quality Dimensions	Elements	Indicators
Prerequisites of quality[11]	**0.1 Legal and institutional environment** – *The legal framework is supportive of statistics.*	0.1.1. The responsibility for compiling statistics is clearly specified. 0.1.2. Data sharing and coordination between data producing agencies is adequate. 0.1.3 Confidentiality of respondents' data is guaranteed and their use is restricted to statistical purposes. 0.1.4 Statistical reporting is ensured through legal mandate and/or measures implemented to encourage voluntary response.
	0.2 Resources - *Resources are commensurate with needs of statistical programs.*	0.2.1 Staff, financial, and computing resources are commensurate with institutional functions. 0.2.2 Measures to ensure the cost-effectiveness of the various statistical programs are implemented.
	0.3. Quality awareness – *Quality is recognized as a cornerstone of statistical work.*	0.3.1 Processes are in place to focus on quality, to monitor the quality of the production and dissemination of statistics, to acknowledge and deal with tradeoffs within quality, and to inform planning.
1. Integrity *Firm adherence to the principle of objectivity in the collection, compilation, and dissemination of statistics.*	**1.1 Professionalism** – *Professionalism in statistical policies and practices is a guiding principle.*	1.1.1 Statistics are compiled on an impartial basis. 1.1.2 Choices of sources and methods are informed solely by statistical considerations. 1.1.3 Statistical agencies are entitled to comment on erroneous interpretation and misuse of statistics.
	1.2 Transparency – *Statistical policies and practices are transparent.*	1.2.1 The terms and conditions under which statistics are produced and disseminated are available to the public. 1.2.2 Internal government access to statistics prior to their release is identified. 1.2.3 Products of statistical agencies/units are clearly identified as such. 1.2.4 Advance notice is given of major changes in methodology, source data, and statistical techniques.
	1.3 Ethical standards – *Statistical processes are guided by ethical standards.*	1.3.1 Guidelines for staff behavior are clear and publicized.

Quality Dimensions	Elements	Indicators
2. Methodological soundness *The conceptual basis for the statistics follows international standards, guidelines, and agreed practices.*	**2.1 Concepts and definitions** – *Concepts and definitions used are in accord with standard statistical frameworks.*	2.1.1 Concepts and definitions: see dataset-specific framework.
	2.2 Scope – *The scope is in accord with internationally accepted standards.*	2.2.1 Scope: see dataset-specific framework.
	2.3 Classification/sectorization – *Classification and sectorization systems are in accord with internationally accepted standards.*	2.3.1 Classification/sectorization systems: see dataset-specific framework.
	2.4 Basis for recording – *Flows and stocks are valued and recorded according to internationally accepted standards.*	2.4.1 Accounting is done on accrual basis. 2.4.2 Market prices are used to value flows and stocks.
3. Accuracy and reliability *Source data and compilation techniques are sound, and disseminated data sufficiently portray reality.*	**3.1 Source data** – *Source data available provide an adequate basis to compile statistics.*	3.1.1 Source data are collected from comprehensive data collection programs that take into account country-specific conditions. 3.1.2 Source data reasonably approximate the definitions, scope, classifications, time of recording, and valuation required. 3.1.3 Source data are timely.
	3.2 Statistical techniques – *Statistical techniques employed conform with sound statistical procedures.*	3.2.1 Data compilation procedures employ sound statistical methods. 3.2.2 Other statistical procedures employ sound statistical methods.
	3.3 Assessment and validation – *Source data are regularly assessed and results validated.*	3.3.1 Source data—including censuses, sample surveys and administrative records—are routinely assessed, e.g., for coverage, sample error, response error, and non-sampling error; the results of the assessments are monitored and made available to inform choices. 3.3.2. Main intermediate results are validated against other information where applicable. 3.3.3 Statistical discrepancies and other potential indicators of problems in statistical outputs are investigated and made available to inform users.

Quality Dimensions	Elements	Indicators
4. Serviceability *Statistics are relevant, timely, consistent, and follow a predictable revisions policy.*	**4.1 Relevance** – *Statistics cover relevant information on the subject field.*	4.1.1 Processes to monitor the relevance and practical utility of existing statistics in meeting users' needs are in place.
	4.2 Timeliness and periodicity – *Timeliness and periodicity follow internationally accepted dissemination standards.*	4.2.1 Periodicity follows dissemination standards. 4.2.2 Timeliness follows dissemination standards.
	4.3 Consistency – *Statistics are consistent over time, internally, and with major data systems.*	4.3.1 Statistics are consistent or reconcilable over a reasonable period of time. 4.3.2 Statistics are internally consistent (e.g., accounting identities observed). 4.3.3 Statistics are consistent or reconcilable with those obtained through other sources and/or statistical frameworks.
	4.4 Revision policy and practice – *Data revisions follow a regular and publicized procedure.*	4.4.1 Revisions follow a regular, well-established and transparent schedule. 4.4.2 Preliminary data are clearly identified. 4.4.3 Studies and analyses of revisions are carried out routinely and made public.

Quality Dimensions	Elements	Indicators
5. Accessibility *Clear data and metadata are easily available and assistance to users is adequate.*	**5.1 Data accessibility** — *Statistics are presented in a clear and understandable manner, forms of dissemination are adequate, and statistics are made available on an impartial basis.*	5.1.1 Statistics are presented in a way that facilitates proper interpretation and meaningful comparisons (layout and clarity of text, tables, and charts). 5.1.2 Dissemination media and formats are adequate. 5.1.3 Statistics are released on a pre-announced schedule. 5.1.4 Statistics are made available to all users at the same time. 5.1.5 Non-published (but non-confidential) sub-aggregates are made available upon request.
	5.2 Metadata accessibility — *Up-to-date and pertinent metadata are made available.*	5.2.1 Documentation on concepts, scope, classifications, basis of recording, data sources, and statistical techniques is available, and differences from international standards are annotated. 5.2.2 Different levels of detail are provided depending on intended audience and type of collection.
	5.3 Assistance to users - *Prompt and knowledgeable support service is available.*	5.3.1 Contact person for each subject field is publicized.

Table 4 Data quality dimensions.

Source: Carson 2001, Toward a Framework for Assessing Data Quality.